

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: May 7, 2004, 11:55:46 ; Search time 50 Seconds  
(without alignments)  
5616.224 Million cell updates/sec

Title: US-09-989-687-4

Perfect score: 890  
Sequence: 1 MFPAAPAPRWLPFLILLLL.....CNKALKPDAKPCESQLCPL 890

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTRMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	3.4	183	11 Q99JPI	Q99JPI mus musculus
2	13	1.5	2165	5 Q19791	Q19791 caenorhabdi
3	12	1.3	833	11 Q8K384	Q8K384 mus musculu
4	12	1.3	845	11 Q8BNJ2	Q8BNJ2 mus musculu
5	11	1.2	129	12 Q8VBA0	Q8VBA0 white spot
6	11	1.2	166	11 Q80Z90	Q80Z90 mus musculu
7	11	1.2	197	11 Q80Z91	Q80Z91 mus musculu
8	11	1.2	230	11 Q80Z92	Q80Z92 mus musculu
9	11	1.2	263	11 Q60819	Q60819 mus musculu
10	11	1.2	406	10 Q8S0L8	Q8S0L8 oryza sativ
11	11	1.2	558	10 Q7XUC6	Q7XUC6 oryza sativ
12	10	1.1	83	3 Q870N2	Q870N2 neurospora
13	10	1.1	143	4 Q15412	Q15412 homo sapien
14	10	1.1	165	4 Q7Z6V1	Q7Z6V1 homo sapien
15	10	1.1	192	6 Q95N24	Q95N24 equus cabal
16	10	1.1	203	11 Q8CH80	Q8CH80 cavia porce

17	10	1.1	245	4 Q7Z502	Q7Z502 homo sapien
18	10	1.1	271	6 Q95KQ0	Q95KQ0 pan paniscu
19	10	1.1	274	6 Q9GKD9	Q9GKD9 bos taurus
20	10	1.1	278	4 Q9BTO9	Q9BTO9 homo sapien
21	10	1.1	292	6 Q9GKE0	Q9GKE0 bos taurus
22	10	1.1	292	6 Q8KMW1	Q8KMW1 bos taurus
23	10	1.1	322	7 Q9GIZ5	Q9GIZ5 homo sapien
24	10	1.1	322	10 Q9S5N5	Q9S5N5 arabidopsis
25	10	1.1	324	7 Q95IB2	Q95IB2 homo sapien
26	10	1.1	324	7 Q95HN6	Q95HN6 homo sapien
27	10	1.1	347	5 Q81209	Q81209 plasmodium
28	10	1.1	413	10 Q84PU6	Q84PU6 oryza sativ
29	10	1.1	414	6 Q8HZR0	Q8HZR0 canis famil
30	10	1.1	439	11 Q8K466	Q8K466 mus musculu
31	10	1.1	439	11 Q8BLH5	Q8BLH5 mus musculu
32	10	1.1	464	13 Q9DGD1	Q9DGD1 gallus gall
33	10	1.1	467	4 Q8NKA4	Q8NKA4 homo sapien
34	10	1.1	467	4 Q9BTY2	Q9BTY2 homo sapien
35	10	1.1	467	4 Q7Z6Y2	Q7Z6Y2 homo sapien
36	10	1.1	481	5 Q18122	Q18122 caenorhabdi
37	10	1.1	492	11 Q7IPW7	Q7IPW7 mus musculu
38	10	1.1	521	11 Q08795	Q08795 mus musculu
39	10	1.1	528	11 Q921X2	Q921X2 mus musculu
40	10	1.1	548	4 Q9BR08	Q9BR08 homo sapien
41	10	1.1	559	11 Q8BRU4	Q8BRU4 mus musculu
42	10	1.1	606	6 Q97554	Q97554 cryctolagus
43	10	1.1	633	6 Q8HZR1	Q8HZR1 canis famil
44	10	1.1	759	6 Q8HZM8	Q8HZM8 equus cabal
45	10	1.1	769	11 Q912V2	Q912V2 rattus norv

## ALIGNMENTS

RESULT 1  
ID Q99JPI PRELIMINARY; PRT; 183 AA.  
AC Q99JPI;  
DT 01-JUN-2001 (TREMBLrel, 17, Created)  
DT 01-JUN-2001 (TREMBLrel, 17, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel, 25, Last annotation update)  
DE Hypothetical protein (Fragment).  
GN ADAMTS8.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_Taxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strauberg R.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC005780; AA05780.1; -  
DR MGD; MGI:1353468; Adamts8.  
DR InterPro; IPR000884; TSP1.  
DR SMART; SM00209; TSP1; 1.  
DR PROSITE; PS50092; TSP1; 1.  
KW Hypothetical protein.  
FT NON\_TER  
SQ SEQUENCE 183 AA; 19960 MW; 5C15HC3E00B3FB3 CRC64;  
Query Match 3.4%; Score 30; DB 11; Length 183;  
Best Local Similarity 100.0%; Pred. No. 7.6e-22;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 741 GQYLLNGNTAISAEIDILVKGTLIKYSGS 770  
Db 34 GQYLLNGNTAISAEIDILVKGTLIKYSGS 63  
RESULT 2  
ID Q19791 PRELIMINARY; PRT; 2165 AA.  
AC Q19791; Q27524;

```

DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE F25H8.3 protein.
GN F25H8.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gajdedy S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copestake T., Cooper C., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Lister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Riffen L., Roopra A., Saunders D., Showkeen R.,
RA Smaldon N., Smith A., Sonhammer E., Straden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sprat J., Wohlman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
DR EMBL; Z69361; CA93288.1; JOINED.
DR EMBL; Z69360; CA93287.1; JOINED.
DR EMBL; Z69360; CA93287.1; JOINED.
DR PIR; T21371; T21371.
DR HSP; P15167; IDTH.
DR MEROPS; M12.135; -.
DR WormPep; F25H8.3; CE05729.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 14.
DR SMART; SM00209; TSP1; 18.
DR PROSITE; PS50215; ADAM_MERPRO; 1.
DR PROSITE; PS50092; TSP1; 15.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 2165 AA; 244397 MW; FCC3DA8AA9C4886 CRC64;

```

Query Match 1.5%; Score 13; DB 5; Length 2165;  
Best local similarity 100.0%; Pred. No. 0.002; Indels 0; Gaps 0;  
Matches 13; Conservative 0; Mismatches 0;

```

QY 536 WGECSTCGGCVQ 548
DB 612 WGECSTCGGCVQ 624

```

RESULT 3

```

ID Q8K384 PRELIMINARY; PRT; 833 AA.
AC Q8K384;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to a disintegrin-like and metalloprotease (reprolysin type)
DE with thrombospondin type 1 motif. 4.
GN ADAMTS4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strauberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027773; AAH27773.1; -.
DR MGD; MGI:133949; Adamts4.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0007223; P:integrin-mediated signaling pathway; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR00884; TSP1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50215; ADAM_MERPRO; 1.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KM Integrin; Protease; Metalloprotease.
SQ SEQUENCE 833 AA; 90097 MW; 4CEC83DFEC3AA619 CRC64;

```

Query Match 1.3%; Score 12; DB 11; Length 833;  
Best local similarity 100.0%; Pred. No. 0.01; Indels 0; Gaps 0;  
Matches 12; Conservative 0; Mismatches 0;

```

QY 539 CSRTCGGCVQPS 550
DB 528 CSRTCGGCVQPS 539

```

RESULT 4

```

ID Q8BNJ2 PRELIMINARY; PRT; 845 AA.
AC Q8BNJ2;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE A disintegrin-like and metalloprotease.
GN ADAMTS4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK083534; BAC38944.1; -.
DR MGD; MGI:133949; Adamts4.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR00884; TSP1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.

```

DR PROSITE; PSS0215; ADAM\_MERPRO; 1.  
 DR PROSITE; PSS0092; TSPI; 1.  
 DR PROSITE; PSS0142; ZINC\_PROTEASE; 1.  
 SQ SEQUENCE 845 AA; 91339 MM; B496C3190D1A9225 CRC64;

Query Match 1.3%; Score 12; DB 11; Length 845;  
 Best Local Similarity 100.0%; Pred. No. 0.01;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 CSRTCGGAGVQFS 550  
 DB 540 CSRTCGGAGVQFS 551

RESULT 5  
 Q8VBA0 PRELIMINARY; PRT; 129 AA.

AC Q8VBA0;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE W5V080 (W5SV137).  
 OS White spot syndrome virus (WSSV).  
 OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae.  
 OX NCBI\_TaxID=92652;

RP SEQUENCE FROM N.A.  
 RA MEDLINE=2154831; PubMed=11699662;  
 RX Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
 RT "Complete genome sequence of the shrimp white spot bacilliform  
 virus.";  
 RL J. Virol. 75:11811-11820(2001).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

RN [3]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=Taiwan;  
 RX MEDLINE=20517548; PubMed=11062040;  
 RA Tsai W.F., Yu H.T., Tzeng H.F., Lwu J.H., Chou C.M., Huang C.J.,  
 Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;  
 RT "Identification and characterization of a shrimp white spot syndrome  
 virus (WSSV) gene that encodes a novel chimeric polypeptide of  
 cellular-type thymidine kinase and thymidylate kinase.";  
 RL Virology 277:100-110(2000).

RN [4]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=Taiwan;  
 RX MEDLINE=21844071; PubMed=11853398;  
 RA Chen L.L., Lei J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,  
 Lo C.F., Kou G.H.;  
 RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white  
 spot syndrome virus and characterization of the motif important for  
 targeting VP35 to the nuclei of transfected insect cells.";  
 RL Virology 293:44-53(2002).

RN [5]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN=Taiwan;  
 RX MEDLINE=21844071; PubMed=11853398;  
 RA Chen L.L., Lei J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,  
 Lo C.F., Kou G.H.;  
 RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white  
 spot syndrome virus and characterization of the motif important for  
 targeting VP35 to the nuclei of transfected insect cells.";  
 RL Virology 293:44-53(2002).

Query Match 1.2%; Score 11; DB 12; Length 129;  
 Best Local Similarity 100.0%; Pred. No. 0.02;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 FLLLLLLLPL 23  
 DB 114 FLLLLLLLPL 124

RESULT 6  
 Q80Z90 PRELIMINARY; PRT; 166 AA.

AC Q80Z90;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Interleukin 15 receptor alpha chain isoform 1C.  
 GN IL15RA.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

RP SEQUENCE FROM N.A.  
 RA STRAIN=C57BL/6;  
 RA Toomey J.A., Gays F., Foster D., Brooks C.G.;  
 RT "Cytokine requirements for the growth and development of mouse NK  
 cells in vitro.";  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 RX EMBL; AY219717; AA062312.1;  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR InterPro; IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam; PF00084; Sush1.1.  
 DR SMART; SM00032; CCP; 1.

QY 14 LLLLLLLLLLPL 24  
 DB 18 LLLLLLLLLLPL 28

Query Match 1.2%; Score 11; DB 11; Length 166;  
 Best Local Similarity 100.0%; Pred. No. 0.026;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 7

Q80Z91 PRELIMINARY; PRT; 197 AA.

AC Q80Z91;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Interleukin 15 receptor alpha chain isoform 1B.  
 GN IL15RA.

OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

RP SEQUENCE FROM N.A.  
 RA STRAIN=C57BL/6;  
 RA Toomey J.A., Gays F., Foster D., Brooks C.G.;  
 RT "Cytokine requirements for the growth and development of mouse NK  
 cells in vitro.";  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 RX EMBL; AY219716; AA062311.1;  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR InterPro; IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam; PF00084; Sush1.1.  
 DR SMART; SM00032; CCP; 1.

QY 14 LLLLLLLLLLPL 24  
 DB 18 LLLLLLLLLLPL 28

Query Match 1.2%; Score 11; DB 11; Length 197;  
 Best Local Similarity 100.0%; Pred. No. 0.03;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## RESULT 8

Q60292 PRELIMINARY; PRT; 230 AA.  
 AC Q60292;  
 DT 01-JUN-2003 (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Interleukin 15 receptor alpha chain isoform 1A.  
 GN IL15RA.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RA Toomey J.A., Gays F., Foster D., Brooks C.G.;  
 RT "Cytokine requirements for the growth and development of mouse NK  
 cells in vitro."  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY219715; AA062310.1; -  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR InterPro; IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam; PF00084; Sush1; 1.  
 DR SMART; SM00032; CCP; 1.  
 DR Receptor.  
 SQ SEQUENCE 230 AA; 24637 MW; 23CFCEAD3AF70F1 CRC64;

Query Match 1.2%; Score 11; DB 11; Length 230;  
 Best Local Similarity 100.0%; Pred. No. 0.034;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLLLPL 24  
 DB 18 LLLLLLLLLLPL 28

## RESULT 9

Q60819 PRELIMINARY; PRT; 263 AA.  
 AC Q60819;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Interleukin 15 receptor precursor.  
 GN IL15RA.  
 OS Mus musculus (Mouse).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D10;  
 RA Gali J.G., Kumaki S., Ahlsten M., Friend D.J., Loomis A., Shanebeck K.,  
 RA Duhose R., Cosman D., Park L.S., Anderson D.M.;  
 RT "Identification and cloning of a novel IL-15 binding protein that is  
 structurally related to the alpha chain of the IL-2 receptor."  
 RL EMBO J. 14:3654-3663 (1995).  
 DR EMBL; U22339; AAC52240.1; -  
 DR PIR; S57346; S57346.  
 DR MGI; MGI:104644; IL15RA.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR InterPro; IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam; PF00084; Sush1; 1.  
 DR SMART; SM00032; CCP; 1.  
 DR Receptor; Signal.  
 KW SIGNAL  
 FT CHAIN 1 32 POTENTIAL.  
 FT INTERLEUKIN 15 RECEPTOR.  
 SQ SEQUENCE 263 AA; 28061 MW; BFCC2CBA5B504 CRC64;

Query Match 1.2%; Score 11; DB 11; Length 263;  
 Best Local Similarity 100.0%; Pred. No. 0.039;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLLLPL 24  
 DB 18 LLLLLLLLLLPL 28

## RESULT 10

Q68018 PRELIMINARY; PRT; 406 AA.  
 AC Q68018;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)  
 DE Putative pectinacetyl esterase.  
 GN B1078G07.15.  
 OS Oryza sativa (Japonica cultivar-group).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC Eriaristidae; Oryzaceae; Oryza.  
 NX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsunoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare (Ga3) genomic DNA, chromosome 1, BAC  
 clone:B1078G07."  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP003407; BAB90194.1; -  
 DR Gramene; Q68018; -  
 DR InterPro; IPR004963; Pec\_acetylase.  
 DR Pfam; PF03283; PAF; 1.4579 MW; C47926770E80A5CB CRC64;  
 SQ SEQUENCE 406 AA; 44579 MW; C47926770E80A5CB CRC64;

Query Match 1.2%; Score 11; DB 10; Length 406;  
 Best Local Similarity 100.0%; Pred. No. 0.057;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 PLLLLLLLLL 22  
 DB 20 PLLLLLLLLL 30

## RESULT 11

Q7XU06 PRELIMINARY; PRT; 558 AA.  
 AC Q7XU06;  
 DT 01-OCT-2003 (TREMBlrel. 25, Created)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE OSJNB001IN17.6 protein.  
 GN OSJNB001IN17.6.  
 OS Oryza sativa (Rice).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC Eriaristidae; Oryzaceae; Oryza.  
 NX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Pu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,  
 RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y.,  
 RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.U., Ding C.W.,  
 RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Han B., Feng Q.,  
 RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,  
 RA Chen L., Fan D.L., Weng Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,  
 RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,  
 RA Qian Y.M., Yang K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,  
 RA Zhang R.Q., Guan J.P., Hong G.F.,  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL606614; CAD41089.1; -  
 SQ SEQUENCE 558 AA; 60905 MW; 87E303311659C93A CRC64;



Query Match 1.1%; Score 11; DB 10; Length 558;  
 Best Local Similarity 100.0%; Pred. No. 0.076;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 15 LLLLLLLLLPLA 25  
 |||||  
 DB 7 LLLLLLLLLPLA 17

## RESULT 12

0870N2 PRELIMINARY; PRT; 83 AA.

AC 0870N2; (TREMBlrel. 24, Created)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Hypothetical protein 49D12.210.  
 GN 49D12.210.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.

OX NCBI\_TaxID=5141;  
 RN [1]

RP SEQUENCE FROM N.A.

RA Schulte U., Aligh V., Hobeisel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatura G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.

RA German Neurospora genome project;  
 RA Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BX95540; CAD79696.1; -  
 KW Hypothetical protein.

SQ SEQUENCE 83 AA; 9293 MW; 8B6B07546FCE6A CRC64;

Query Match 1.1%; Score 10; DB 3; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 0.15;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 14 LLLLLLLLLLP 23  
 |||||  
 DB 18 LLLLLLLLLLP 27

## RESULT 13

015412 PRELIMINARY; PRT; 143 AA.

AC 015412; (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last annotation update)  
 DE CTG4A.  
 GN CTG4A.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=97369492; PubMed=9225980;  
 RA Margolis R.L., Abraham M.R., Gatchell S.B., Li S.H., Kidwai A.S.,  
 RA Breschel T.S., Stine O.C., Callahan C., McInnis M.G., Rose C.A.;  
 RT "CDNs with long CAG trinucleotide repeats from human brain.";  
 DR HUM. Genet. 100:114-122(1997).

DR EMBL; U08074; AAB91442.1; -  
 SQ SEQUENCE 143 AA; 15959 MW; 0F2119BFA33A1E6 CRC64;

Query Match 1.1%; Score 10; DB 4; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 0.24;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 14 LLLLLLLLLLP 23  
 |||||  
 DB 17 LLLLLLLLLLP 26

## RESULT 14

0726V1 PRELIMINARY; PRT; 165 AA.

AC 0726V1; (TREMBlrel. 25, Created)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE DJ20N2.5.2 (Novel protein (MG1314) similar to fucosidase, alpha-L-1,  
 DE tissue, variant 2).  
 GN DJ20N2.5.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.

RA Kay M.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL031320; CAD92485.1; -  
 SQ SEQUENCE 165 AA; 19835 MW; 0427DB86CD215820 CRC64;

Query Match 1.1%; Score 10; DB 4; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 0.27;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 14 LLLLLLLLLLP 23  
 |||||  
 DB 13 LLLLLLLLLLP 22

## RESULT 15

095N24 PRELIMINARY; PRT; 192 AA.

AC 095N24; (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Aggrecaase-1 (Fragment).  
 OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OX NCBI\_TaxID=9796;  
 RN [1]

RP SEQUENCE FROM N.A.

RA Flannery C.R., Little C.B.;  
 RL "Expression and activity of equine aggrecaases";  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF368321; AAK53425.1; -  
 DR MEROPS; M12.221; -  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.

DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.

DR PROSITE; PS0215; ADAM\_MEROP; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.

FT NON\_TER 1  
 FT NON\_TER 192

SQ SEQUENCE 192 AA; 20670 MW; 9013B0E19FCE8C56 CRC64;

Query Match 1.1%; Score 10; DB 6; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 0.31;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 495 PMADGTPGCP 504  
 |||||  
 DB 157 PMADGTPGCP 166

Fri May 7 12:18:01 2004

us-09-989-687-4.rpt

Page 6

Search completed: May 7, 2004, 12:00:09  
Job time : 52 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: May 7, 2004, 11:43:25 ; Search time 18 seconds  
(without alignments)  
2574.580 Million cell updates/sec

Title: US-09-989-687-4  
Perfect score: 4853  
Sequence: 1 MFPAAPAPRMPLFLLILL.....CNKALKEPAKPCESQICPL 890

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	4853	100.0	890 1	AT88_HUMAN
2	3918.5	80.7	905 1	AT88_MOUSE
3	2344	48.3	968 1	AT81_MOUSE
4	2343.5	48.3	967 1	AT81_MOUSE
5	2328.5	48.0	967 1	AT81_MOUSE
6	2138.5	44.1	950 1	AT81_MOUSE
7	1842.5	38.0	837 1	AT81_MOUSE
8	1768.5	36.5	930 1	AT85_MOUSE
9	1734.5	35.7	930 1	AT85_MOUSE
10	1720.5	35.5	1906 1	AT80_MOUSE
11	1699.5	35.0	1935 1	AT89_MOUSE
12	1644.5	33.9	630 1	AT84_MOUSE
13	1633	33.6	1911 1	AT80_MOUSE
14	1284.5	26.5	562 1	AT85_MOUSE
15	1200	24.7	997 1	AT87_MOUSE
16	1189	24.5	1593 1	AT82_MOUSE
17	1124.5	23.2	1081 1	AT88_MOUSE
18	1108	22.8	1077 1	AT80_MOUSE
19	1108	22.8	1211 1	AT82_MOUSE
20	1102	22.7	1205 1	AT82_MOUSE
21	1072.5	22.1	1213 1	AT82_MOUSE
22	1070	22.0	1224 1	AT86_MOUSE
23	1049	21.6	1205 1	AT83_MOUSE
24	1016.5	20.9	1223 1	AT84_MOUSE
25	1004	20.7	1095 1	AT87_MOUSE
26	989	20.4	1210 1	AT89_MOUSE
27	979	20.2	860 1	AT86_MOUSE
28	971	20.0	1207 1	AT89_MOUSE
29	615.5	12.7	245 1	AT84_MOUSE
30	538	11.1	207 1	AT85_MOUSE
31	498	10.3	377 1	AT81_MOUSE
32	488.5	10.1	525 1	AT81_MOUSE
33	428.5	8.8	824 1	AD08_MOUSE

34	344.5	7.1	920 1	AD19_MOUSE	035674 mus musculus
35	341.5	7.0	776 1	AD28_MOUSE	095816 macaca fasc
36	336.5	6.9	813 1	AD33_MOUSE	095211 homo sapien
37	332.5	6.9	956 1	AD19_MOUSE	095013 homo sapien
38	330.5	6.8	793 1	AD28_MOUSE	095116 mus musculus
39	325	6.7	819 1	AD09_MOUSE	013443 homo sapien
40	323	6.7	775 1	AD28_MOUSE	095042 homo sapien
41	318	6.6	857 1	AD22_MOUSE	095116 mus musculus
42	317.5	6.5	754 1	AD07_MOUSE	095219 homo sapien
43	314	6.5	864 1	AD15_MOUSE	088839 mus musculus
44	311	6.4	909 1	AD12_MOUSE	043184 homo sapien
45	309.5	6.4	826 1	AD08_MOUSE	005910 mus musculus

## ALIGNMENTS

RESULT 1  
ID AT88\_HUMAN STANDARD; PRT; 890 AA.  
AC Q9UP79; Q9NZS0;  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2) (METH-8).  
GN ADAMTS8 OR METH2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=99367466; PubMed=10438512;  
RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S., Lombardo M., Iruela-Arispe M.L.;  
RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new family of proteins with angio-inhibitory activity.";  
RL J. Biol. Chem. 274:23349-23357(1999).  
RN [2]  
RP SEQUENCE OF 195-440 FROM N.A.  
RX MEDLINE=20079168; PubMed=10610729;  
RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;  
RT "ADAM-TS8, a novel metalloproteinase of the ADAM-TS family located on mouse chromosome 9 and human chromosome 11.";  
RL Genomics 62:312-315(1999).  
CC - FUNCTION: Has anti-angiogenic properties.  
CC - COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC - SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).  
CC - TISSUE SPECIFICITY: Highly expressed in adult and fetal lung, lower expression in brain, placenta, heart, stomach and fetal brain and kidney.  
CC - DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.  
CC - PTM: The precursor is cleaved by a furin endopeptidase (By similarity).  
CC - SIMILARITY: Belongs to peptidase family M12B.  
CC - SIMILARITY: Contains 1 disintegrin-like domain.  
CC - SIMILARITY: Contains 2 TSP type-1 domains.  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC EMBL: AF060153; AAD48081.1; -;  
CC EMBL: AF175283; AAF25806.1; -;

DR HSSP; P34179; 11AG.  
 DR MEROPS; M12.226; -.  
 DR Genew; HGNC:224; ADAMTS8.  
 DR MIM; 605175; -.  
 DR GO; GO:0005178; F: integrin binding; TAS.  
 DR GO; GO:0008237; F: metalloproteinase activity; TAS.  
 DR GO; GO:0007345; F: embryogenesis and morphogenesis; TAS.  
 DR GO; GO:0008285; P: negative regulation of cell proliferation; TAS.  
 DR InterPro; IPR001762; D: integrin.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF00421; Repolysin; 1.  
 DR Pfam; PF00090; TSP1; 2.  
 DR PRINTS; PRO1705; TSP1REPEAT.  
 DR SMART; SM00209; TSP1; 2.  
 DR PROSITE; PS50215; ADAM\_MEROPS; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PS50092; TSP1\_2.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR HydroLase; Metalloprotease; Zinc; signal; Glycoprotein; Zymogen;  
 Repeat; Extracellular matrix; Heparin-binding.  
 KM SIGNAL 1 27  
 FT PROPEP 28 214  
 FT CHAIN 25 890  
 FT DOMAIN 215 430  
 FT DOMAIN 439 526  
 FT DOMAIN 527 582  
 FT DOMAIN 584 690  
 FT DOMAIN 691 832  
 FT DOMAIN 834 889  
 FT DOMAIN 890 905  
 FT METAL 364 364  
 FT ACT SITE 365 365  
 FT METAL 368 368  
 FT METAL 374 374  
 FT CARBOHYD 345 345  
 FT CARBOHYD 401 401  
 FT CARBOHYD 466 466  
 FT CARBOHYD 491 491  
 FT CARBOHYD 600 600  
 FT CONFLICT 195 195  
 FT CONFLICT 413 440  
 SQ SEQUENCE 890 AA; 96671 MW; 57D0EB03D5739D3 CRC64;  
 Query Match 100.0%; Score 4853; DB 1; Length 890;  
 Best Local Similarity 100.0%; Pred. No. 2e-293;  
 Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 1 MFPPAPAPRLPFLILLILLIPLAGAPAPAPAGQASLVPPTLPSSAGELALHISA 60  
 DB 1 MFPPAPAPRLPFLILLILLIPLAGAPAPAPAGQASLVPPTLPSSAGELALHISA 60  
 DB 61 FGKFFVRLAPDDSLAPEFKIRLGSGGATGEGRLGCGFSGTVNGEPESLAAVSIC 120  
 DB 61 FGKFFVRLAPDDSLAPEFKIRLGSGGATGEGRLGCGFSGTVNGEPESLAAVSIC 120  
 DB 121 RGLSGSLDGEETITPOGAGSLAOPHRLQWGPAGAPLPFRGPEWEVTEGEQORER 180  
 DB 121 RGLSGSLDGEETITPOGAGSLAOPHRLQWGPAGAPLPFRGPEWEVTEGEQORER 180  
 DB 121 RGLSGSLDGEETITPOGAGSLAOPHRLQWGPAGAPLPFRGPEWEVTEGEQORER 180  
 DB 181 GDHDEDESESESEAGSEPPPLGATSTRKFFVSEAFVETLLVADASNAAFYADL 240  
 DB 181 GDHDEDESESESEAGSEPPPLGATSTRKFFVSEAFVETLLVADASNAAFYADL 240  
 DB 241 QNHLLTMSVAAARYKPKSTKSNLMVYVLLVEDEKGPPEVSDNGGLTKRNCWQOR 300  
 DB 241 QNHLLTMSVAAARYKPKSTKSNLMVYVLLVEDEKGPPEVSDNGGLTKRNCWQOR 300

QY 301 FNQSDRPHEDYDAILLITRONFCGQEGLCPTLGVADIGTICDPKSCSVIIEDEGLQAH 360  
 DB 301 FNQSDRPHEDYDAILLITRONFCGQEGLCPTLGVADIGTICDPKSCSVIIEDEGLQAH 360  
 QY 361 TLALHELGHVLSMHPHDSKCTFLFGPMKGHVMAFLFHLNQTLPMSGSMYITELLDG 420  
 DB 361 TLALHELGHVLSMHPHDSKCTFLFGPMKGHVMAFLFHLNQTLPMSGSMYITELLDG 420  
 QY 421 GHSDCLLDAPGALPLPTGLPGRVALYQJDOQCEQIFGPPFRHCENTSADQVCAQLMCHT 480  
 DB 421 GHSDCLLDAPGALPLPTGLPGRVALYQJDOQCEQIFGPPFRHCENTSADQVCAQLMCHT 480  
 QY 481 DGAEPLCHTKNSLPMADSTPCGPHLCSGSCLEEEVERPKFVDDGMAPWGWGBCS 540  
 DB 481 DGAEPLCHTKNSLPMADSTPCGPHLCSGSCLEEEVERPKFVDDGMAPWGWGBCS 540  
 QY 541 RTGGGVQFSHRECKDPFONGARYCLGRANKYQSCHEECPPDGKSFREOCCERYANN 600  
 DB 541 RTGGGVQFSHRECKDPFONGARYCLGRANKYQSCHEECPPDGKSFREOCCERYANN 600  
 QY 601 YTMDSNLLQWPVKYAGVSPRCKLPFRARSEFKVFEAKVIDGTLGPEETALICRG 660  
 DB 601 YTMDSNLLQWPVKYAGVSPRCKLPFRARSEFKVFEAKVIDGTLGPEETALICRG 660  
 QY 661 QCVKACDHYVSPKRLDKGVCGGKNSCRKVSGLTPTVGYNDITIPAGATNIDVK 720  
 DB 661 QCVKACDHYVSPKRLDKGVCGGKNSCRKVSGLTPTVGYNDITIPAGATNIDVK 720  
 QY 721 QRHREVNQDGYLAKTADGQYLNLGALSAIQDILVYGTILKSGSATLERLOSF 780  
 DB 721 QRHREVNQDGYLAKTADGQYLNLGALSAIQDILVYGTILKSGSATLERLOSF 780  
 QY 781 RPLREPLVQLITVPEVPPRYKTYFPVNDVDFSMQSKERATNIIQPLLAQWYLG 840  
 DB 781 RPLREPLVQLITVPEVPPRYKTYFPVNDVDFSMQSKERATNIIQPLLAQWYLG 840  
 QY 841 DWEGCSGTGAGMORTVECPDPSGASATCKALKPEADPCSOQL 890  
 DB 841 DWEGCSGTGAGMORTVECPDPSGASATCKALKPEADPCSOQL 890  
 RESULT 2  
 ID ADAMTS8 MOUSE STANDARD; PRT: 905 AA.  
 AC PRT110;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2).  
 GN ADAMTS8.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20079168; PubMed=10610729;  
 RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;  
 RT "ADAM-TS8, a novel metalloproteinase of the ADAM-TS family located on  
 RT mouse chromosome 9 and human chromosome 11.";  
 RL Genomics 62:312-315(1999).  
 CC -1- FUNCTION: Has anti-angiogenic properties (By similarity).  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (By similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed specifically in adult lung and heart  
 CC and low expression during mouse development.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix. (By  
 CC -1- PVM: The precursor is cleaved by a furin endopeptidase (By  
 CC similarity).

CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 2 TSP type-1 domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR HSSP: AFI75282; AAF5805.1; -.  
 DR HSSP: P34179; 1IAC.  
 DR MEROPS: M12.226; -.  
 DR WGD: MGI1353468; Adamts8.  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR006025; Pept\_M Zn BS.  
 DR InterPro: IPR001590; Peptidase\_M12b.  
 DR InterPro: IPR000884; TSP1.  
 DR InterPro: IPR008085; TSP1.  
 DR Pfam: PF01421; Reprolysin; 1.  
 DR Pfam: PF00090; TSP\_1; 2.  
 DR PRINTS: PRO1705; TSP1REPEAT.  
 DR SMART: SMO0209; TSP1; 2.  
 DR PROSITE: PSS0215; ADAM\_MEROPS; 1.  
 DR PROSITE: PSS0437; DISINTEGRIN\_1; FALSE NEG.  
 DR PROSITE: PSS0214; DISINTEGRIN\_2; FALSE NEG.  
 DR PROSITE: PSS0092; TSP1; 2.  
 DR PROSITE: PSS0142; ZINC\_PROTEASE; 1.  
 DR HydroLase: Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KM Repeat; Extracellular matrix; Heparin-binding.  
 FT SIGNAL 1 28  
 FT PROPEP 29 228  
 FT CHAIN 229 905  
 FT DOMAIN 229 452  
 FT DOMAIN 453 541  
 FT DOMAIN 542 597  
 FT DOMAIN 599 705  
 FT DOMAIN 706 847  
 FT DOMAIN 848 904  
 FT METAL 378 378  
 FT ACT\_SITE 379 379  
 FT METAL 382 382  
 FT METAL 388 388  
 FT CARBOHYD 415 415  
 FT CARBOHYD 480 480  
 FT CARBOHYD 506 506  
 FT CARBOHYD 615 615  
 SQ SEQUENCE 905 AA; 96879 MW; 124D4132B3A0CAE CRC64;  
 Query Match 80.7%; Score 3918.5; DB 1; Length 905;  
 Best Local Similarity 79.2%; Pred. No. 1.7e-235;  
 Matches 727; Conservative 74; Mismatches 99; Indels 15; Gaps 4;  
 QY 1 MFPAAPAPRLPPLILLILL-LLPLARGAPAPPAAGGSELVPRPLPGSGELAHIS 59  
 DB 1 MURDPTTGMPLLLILLQPPPLVCGAPAGGTAAQASSELVPRPLPGSASLEAFHLS 60  
 QY 60 AFKSGFVLRLAPDSFLAFEFKIERLGSGRATGGRGGRGFFSGTVNGEESIAAVSL 119  
 DB 61 AFGQGFVLRLAPDSFLAFEFKIERLGSAAAGSPGLRGCFSGTVNGEESIAAVSC 120  
 QY 120 CRGLSGSFLLDGESEFTIQPGAGSGLAOPHRLQRMKPAAR-----PLRGPWM 168  
 DB 121 VAWMSGSFLLAGSEFTIQPGAGSGLDQPHRLQRMKPGRRDPGLAAAEVPLDQLEW 180  
 QY 169 EVELTGSGQROERDCHODESEESQE--EBAEGASEPPPLGATSRTRFVSEARFVETLL 226  
 DB 181 EVMGNGQGERSDNEDRKQDEKGLKETEDSRKXPPPPGSKTRKRVSEARFVETLL 240  
 QY 227 VADASMAAFYGDLOHNLITLMSVAAIRYKHPISIKSIVLAVYKVLIVEDKMGPEVSDN 286

DB 241 VADASMAAFYGTDIQNLHILTMSPMAIRYKHPISIRNSVNLVYKVLIVKEMKPEVSDN 300  
 QY 287 GGLTLRNFECNQRREPNQSDRHPHYDTAILLTRNFCGSGGLCDLIGVADIGTICDPNK 346  
 DB 301 GGLTLRNFCSQRRFPNPSDRHPHYDTAILFTRNQFCGSGGLCDLIGVADIGTICDPNK 360  
 QY 347 SCSTVEDGGLQAAHTLHGLHYLSMHPDSEKPTPLRFGPKRHHWAPLFLVNLQTLPW 406  
 DB 361 SCSTYKDEGLQAAHTLHGLHYLSMHPDSEKPCVRLFGPKRHHWAPLFLVNLQTLPW 420  
 QY 407 SPCAMVTELLDGGHGDCLIDAPGALPLPTGPGSMALYQLDQCRQITFGPDPFCEN 466  
 DB 421 SPCAVVTELLDGGHGDCLIDAPGALPLPTGPGSTLYELDQCKQITFGPDPFCEN 480  
 QY 467 TSAQDVCAQMC-HTDGEPLCHTKNGSLPMADGTPCGPHLCSGSCLPBEVEPRKV 525  
 DB 481 TSVEDICVQLCARRDDEPTCHTKNGSLPMADGTPCGPHLCLDGSCLVLEVENPKAV 540  
 QY 526 VDGMPAPGMPWGESRFGCGGVSPSHRECKDPPEPONGRXYCLGRARYQSCHTCECPDG 585  
 DB 541 VDGMPAPGMPWGESRFGCGGVSPSHRECKDPPEPONGRXYCLGRARYQSCHTCECPDG 600  
 QY 586 KSFEQCEKXNAYNTDMDGNLLQWPKYAGVSPDRCKLFCRAGRSEKVEEAYID 645  
 DB 601 KSFEQCEKXNAYNTDMDGNLLQWPKYAGVSPDRCKLFCRAGRSEKVEEAYID 660  
 QY 646 GTLGGPEPLACVAGQCVKAGCDHVNSPKLDCGVCGGKNGSCRYVSGSLPTNYGN 705  
 DB 661 GTLGGPEPLACVAGQCVKAGCDHVNSPKLDCGVCGGKNGSCRYVSGSLPTNYGN 720  
 QY 706 DIVTIPAGATNIDVQKSHRPGVNDGNVYLLKTDAGQYLLNGNLASIAIEDILVKGITL 765  
 DB 721 DIVTIPAGATNIDVQKSHRPGVNDGNVYLLKTDAGQYLLNGNLASIAIEDILVKGITL 780  
 QY 766 KYSSSIATLELQSFRLPEPLTYQLITVPEGEVPPKXKTTFFPNPNDVDSMOSKERAT 825  
 DB 781 KYSSSMATLELQSFRLPEPLTYQLITVPEGEVPPKXKTTFFPNPNDVDSMOSKERAT 840  
 QY 826 TNIQPLHMQVYVGDSESGSTCGAMORRTVECRPDSQASATCKNAKLPEDAKPES 885  
 DB 841 TNIQPLHMQVYVGDSESGSTCGAMORRTVECRPDSQASATCKNAKLPEDAKPES 900  
 QY 886 QLCPL 890  
 DB 901 QLCPL 905  
 RESULT 3  
 ATSL\_MOUSE STANDARD; PRT; 968 AA.  
 ID ATSL\_MOUSE  
 AC P97857; OS4768;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).  
 GN ADAMTS1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=128/SyJ;  
 RX MEDLINE=96110583; PubMed=9441751;  
 RA Kuno K., Liza H., Ohno S., Matsushima K.;  
 RT "The exon/intron organization and chromosomal mapping of the mouse  
 RL ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";   
 RV Genomics 46:466-471(1997).  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97150761; PubMed=8995297;  
 RA Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,



```

QY 124 SSGFLDGEFTTPOGAGSLAQPHRLQWGPAGAPPLPRGPEMEVETGEG-----QROE 179
DB 153 RGAFFYLQGEFFFIQAP-----GVATRLAPNAYEESSARPOPHILRRR 198
QY 180 RGP-----HOEDSEESQDEEAGSEPPPLGA-----TSRTKQPVSEA 219
DB 199 RSGGAKCGVWDEDELPTSDRPSQNTNQPWVRPDPQDAGKPSGSIKKRFVSSP 258
QY 220 RFVETLLVADAAFAFGADLOHILTLMSVARIYKHSINKSNILMYKXLYIYEDDEM 279
DB 259 RYVETMLVADQSMADPHSGSLKXLYLTLVSAARFKHSINKSLVYVKLVYEBK 318
QY 280 GPEVSNNGITLTFNFCNMQRRNPQSDRPHRYDTAILLTRONFCQSEGLCDTLGVADIG 339
DB 319 GPEVTNNAITLTFNFCNMQRRNPQSDRPHRYDTAILLTRONFCQSEGLCDTLGVADIG 377
QY 340 TIGDPKSGSVIYDEGLQAAHTLAHLGHVLSMHPHDSKPCRRLLGPMKHHVMAPLFVH 399
DB 378 TVDPSRSGSVIYDEGLQAAHTLAHLGHVLSMHPHDSKPCRRLLGPMKHHVMAPLFVH 437
QY 400 LMGTLFWSPCSAMYTELLDGHGDCLLDAPGALPLPTLPGRMALYOLDQCRQITGP 459
DB 438 LHSQWSPSCSAVMYVSLFDNGHGECLMDXPNPILPESDLG--TLVDANNQCGFTTGE 495
QY 460 DPHRCNTSAQDVCAQLNC-HTDGAAPLCHTKNGSLPMDGTPCGPRLCSGSCLPBEB 518
DB 496 ESKHCPD--AASCTTLMCTGSGLLVQYQYH--FPMADGTSCEGKVCVSGKCVNTXD 551
QY 519 VERPKFVNDGMAFPMGECSTCGGYQFHSRECKDEPEONGRGVCLGRKATYQSCHT 578
DB 552 MHPATFVNGSNCPMPMGDCSTCGGYQYMEEDNVPKNGKCYGKRVIRSCNI 611
QY 579 EBCPP-DGSGFPEQCEKTNAYNTYDMGN--LQWPKYAGVSPRCKLFCRARGSE 635
DB 612 EDCPDNNGKTFREEQCEANHEFSKSF--GNEPTEVETPKYAGVSPRCKLFCRARGSE 670
QY 636 FVYFEAKVLDGTLGCEETLALICVROGCVKAGDHVDSRKLDKGCVGCKNSGRKYSG 695
DB 671 FVYLDKAVDGTFCPSDSTSVCCVQGCYVAGCDRIIDSKKDKGCVGCKNSGRKYSG 730
QY 696 SLPTNYGVNDVITIPAGATNIDVKORSHPGVQNDGNYLALXTADGOYLLNGLNLSAIE 755
DB 731 IYVSTRPGVHDVITIPAGATNIEVKRNRGRSRRNGSFLAIPADGTYILNGLNLSLIE 790
QY 756 QDLVYGTILTKYSGSIALTERLQSPRLPEPLTVQLTYPGVFPKXKTYTFEVNDVDF 815
DB 791 QDLTYGTILTKYSGSIALTERLQSPRLPEPLTVQLTYPGVFPKXKTYTFEVNDVDF 843
QY 816 SMOSKERATVNIICPLHAQWVLGDMWSECSSTCGAMORRTVECRDPSGASATCNKAL 875
DB 844 ----KKTESFPAIPIPF--SEWVIEBWGSCSTKCSGMRVYQCRDINGHNASCAKEV 897
QY 876 KPEDAKPCESQUCP 889
DB 898 KPASTRPCADLPCP 911

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-Brain;
RA Liu X., Yu Y., Yin T., Johnstone E.W., Stephenson D.T., Clemens J.A.,
RA Little S.P.;
RT "Induction of a disintegrin and metalloprotease with the
RT chromospondin type I motif (ADAMTS).";
RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 18-967 FROM N.A.
RC STRAIN-Sprague-Dawley; TISSUE-Liver;
RX MEDLINE=20304099; PubMed=10847486;
RA Diamantis I., Luetth M., Hoevel M., Reichen J.;
RT "Cloning of the rat ADAMTS-1 gene and its down regulation in
RT endothelial cells in cirrhotic rats.";
RL Liver 20:165-172(2000).
CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
CC involved in its turnover. Has angiogenic inhibitor activity (by
CC similarity). Active metalloprotease, which may be associated with
CC various inflammatory processes as well as development of cancer
CC cachexia. May play a critical role in follicular rupture (by
CC similarity).
CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1683-Glu-Leu-1684
CC site, within the chondroitin sulfate attachment domain.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -1- INDUCTION: Down-regulated in endothelial cells derived from
CC cirrhotic liver.
CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -1- SIMILARITY: Belongs to peptidase family M12B.
CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF149118; AAC34012.1; -
DR EMBL; AF304446; AAC29823.1; -
DR MEROPS; M12.222; -
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006025; Pept_M12B_BS.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00030; tsp_1; 3.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS50215; ADAM_MEROP; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS50093; TSP1; 3.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL 1 54
FT PROPEP 55 252
FT CHAIN 253 967
FT DOMAIN 253 475
FT DOMAIN 476 558
FT DOMAIN 559 614
FT TSP TYPE-1 1.

```

```

FT DOMAIN 616 724 CYS-RICH.
FT DOMAIN 725 857 SPACER.
FT DOMAIN 854 910 TSP TYPE-1.2.
FT DOMAIN 911 967 TSP TYPE-1.3.
FT DOMAIN 194 198 POLY-ARG.
FT SITE 205 205 CYSTEINE SWITCH (POTENTIAL).
FT METAL 401 401 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT SITE 402 402 BY SIMILARITY.
FT METAL 405 405 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 411 411 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 764 764 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 782 782 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 945 945 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 21 21 I -> V (IN REF. 2).
FT CONFLICT 26 31 KFRSSQ -> RSRGSL (IN REF. 2).
FT CONFLICT 49 49 V -> A (IN REF. 2).
FT CONFLICT 72 72 R -> P (IN REF. 2).
FT CONFLICT 79 79 L -> G (IN REF. 2).
FT CONFLICT 249 249 R -> G (IN REF. 2).
FT CONFLICT 262 265 TMLV -> NLK (IN REF. 2).
FT CONFLICT 607 607 S -> F (IN REF. 2).
FT CONFLICT 936 936 L -> V (IN REF. 2).
FT CONFLICT 962 962 I -> T (IN REF. 2).
SQ SEQUENCE 967 AA; 105705 MW; F93C864F6DCDB4CF CRC64;

Query Match 48.3%; Score 2343.5; DB 1; Length 967;
Best Local Similarity 49.5%; Pred. No. 8e-138;
Matches 451; Conservative 143; Mismatches 243; Indels 75; Gaps 19;

14 LLLLLLLLLPLARAPARPAAGASLWPT--RLPGSAGELALHLAFGKGVRLAP 71
38 LLLASITMLCTCRAGHGRPT--EDBLVLPSELRAGHSTLLRLDAFGQDHLKLP 95
72 DDSFLADEFKIERLGSGRATGER-----GIRGCFSGTYNGEPESILAVSLCRGLS 124
96 DSGFLAGFTLQTV---GRSPGEAQLDPTGLAHCFYGTGNDPSSAALSLCEGVR 152
125 GSFLLDDEFTIQOQAG-----SLAQPH---RLQRMQPARARLPFGPM 168
153 GAFYLOGEEFTIQAPAVATERLVPAPKEISLAPPEFHILRRRRSGGAKC----- 205
169 EVETGEGORRGHODESEESQEEAEAGSEPPPLGA-----TSRTKAFVSEAR 221
206 -----GVMDEETLPTNSGREGSNTDPQWPLNRPFTQAGKPTGSGIKKAFVSPPY 259
222 VETLLVADASAAAYGADLQNHITLMSVAARITKHSIKSINLMVYKLVYEDKMG 281
260 VETMLVADQSMADHGGSLKXVLLTLFVSVAAPFKHSHINSISLVVVKLVYEQKP 319
282 EVSNGGGLTLNFCNMORRFNPSDRPEHYDTLILTRONFCQEGICDILGVADIGT 341
320 EYTNALTLNPFCSKQKSHSPDRPEHYDTLILTRDQLGSH--TCDTLMDADGTV 378
342 CDPKSCSVIEDEGLQAATLAHELGHVLSWPHDSKPCRLRFPMKHHVVAFLPHLN 401
379 CDPKSCSVIEDDGLQAATLAHELGHVFNPHDAGKCSFNGVSDSHLMSMSLSD 438
402 QTLMSPCSAMYLIELLDGGHGLLAPGAALPLPLGLGGRMALYOLDOOQOIFGDP 461
439 HSQSPSPSAYMVSFLNNGHGBCLMDKPNPLPDLPLG--TLVYANQCOQTFRES 496
462 RHCPNTSAQVCAQOLMC--HTDGAAPLCTKNKSLPMADGTPCGPHLCSGSGCLPEEVE 520
497 THCPD--AATCSTLMCTGTGGGLVQCCKA--FPMADGISCGRKCVSGKCNKTKDMK 552
551 RPKPVVDGMAFMGWSGCSRTCCGGVGFSHRBECKDEPKNGGGRVYCLGRAYOSCTEE 580
553 HPATPVHSGWGMGWSGCSRTCCGGVGVYTRBECDPPYKNGKYGCKAKVRYNSCTIED 612
581 CPP-DGKSFREQCCEKYNAYNTMDGN--LLQWPKYAGVSPDRCKLFCRARGSEFK 637

```

```

DB 613 CPDNNKTFREEOCEAHNEFSKSF--GNEPTWETPKYAGVSPDRCKLTCCKXGICGYF 671
QY 638 VFEAKVIDGTLCGETLAIQVRGQCVAGCDHVVDSPKLDKCGVCGKXNSGRKXGSL 697
DB 672 VLQKRVYDGTFCSPDSSTVCQCGCVAKGCDRIIDSKKFKDGCVCAGNSTCKKISGTV 731
QY 698 TPTNYGNDIYTTAGATNIDYKORSHPGVQNDGNATLAKTADQVYLLNGLNLSAIED 757
DB 732 TSTPGYNDIYTTAGATNIEVKHNPGRSGKNSFLAIRADQFTYLLNGFTLSTLEOD 791
QY 758 IIVGTLIKYGSATLERLQSFRLPEPLTVQLITVGEVFPKXTEFEVPNDVDFSM 817
DB 792 LTYGQTVARYGSSAALERIRSFSLKEPLITQVLNV--GHARPKIKYTFM----- 842
QY 818 QSSKERATTNIQPLAHQVYLGWSGCSSTCCGAGMQRRTVEGCRDPGQASATCNKALKP 877
DB 843 --KKTEBFNAPIPF--SEWVIEEMGCSKTCGGWQRVVECDINGHPASBCKAEVKP 898
QY 878 EDAPCESSQLCP 889
DB 899 ASTRPCADLPCEP 910

RESULT 5
ID ATSL HUMAN STANDARD; PRT; 967 AA.
AC Q9UHI8; Q9NSJ8; Q9P2K0; Q9UH83; Q9UP80;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1) (METH-1).
GN ADAMTS1 OR METH1 OR KIAA1346.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Casas C., Pritchard M.A., Estivill X., Arbones M.L.;
RT "Cloning, characterization and mapping on human chromosome 21 of the
RT orthologue of murine Adams-1."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND FUNCTION.
RA Tissue=Heart;
RC MEDLINE=99367466; PubMed=10438512;
RX Valquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,
RA Lombard M., Inuela-Arispe M.L.;
RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new
RT family of proteins with angio-inhibitory activity."
RL J. Biol. Chem. 274:23349-23357(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Tissue=Endothelial cells;
RC MEDLINE=20247184; PubMed=10785405;
RX Gienke J., Schmitt A.O., Pilarsky C., Hinzmann B., Weiss B.,
RA Rosenthal A., Therauch K.H.;
RT "Differential gene expression by endothelial cells in distinct
RT angiogenic states."
RL Eur. J. Biochem. 267:2820-2830(2000).
RN [4]
RP SEQUENCE FROM N.A.
RA Tissue=Brain;
RC MEDLINE=20181126; PubMed=10718198;
RX Nagata T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 7:65-73(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;

```



RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,  
 RA Park H.-S., Toyoda A., Ishii K., Totsuki Y., Choi D.-K., Gonser Y.,  
 RA Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,  
 RA Polley A., Menzel U., DeLabar T., Kumpf K., Lehmann R., Patterson D.,  
 RA Reichwald K., Kump A., Schillhabel W., Schudy A., Zimmermann W.,  
 RA Rosenthal A., Kudoh T., Shibuya K., Kawasaki K., Antonarakis S.E.,  
 RA Shintani A., Sasaki T., Nagamine K., Mitsuura S., Antonarakis S.E.,  
 RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,  
 RA Schaefer M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecher H.,  
 RA Ramser J., Beck A., Klages S., Hennig S., Resselmann L., Dagand E.,  
 RA Mehner J., Borzym K., Gardiner K., Nizetic D., Francis F.,  
 RA Lehnach H., Reinhardt K., Yaspo M.-L.,  
 RT "The DNA sequence of human chromosome 21.",  
 RL Nature 405:311-319(2000).  
 RN [6]  
 RP SEQUENCE OF 418-967 FROM N.A.  
 RC TISSUE=Melanoma;  
 RA Blum H., Baurerachs S., Mewes H.-W., Weil B., Wiemann S.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
 CC involved in its turnover (By similarity). Has angiogenic inhibitor  
 CC activity. Active metalloprotease, which may be associated with  
 CC various inflammatory processes as well as development of cancer  
 CC cachexia. May play a critical role in follicular rupture.  
 CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1938-Glu-1-Leu-1939  
 CC site, within the chondroitin sulfate attachment domain.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (By similarity).  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix.  
 CC -1- PFM: The precursor is cleaved by a furin endopeptidase (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 3 TSP type-1 domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF170084; AAF15317.1; -  
 DR EMBL: AF060152; AAD48080.1; ALT\_INTT.  
 DR EMBL: AF207664; AAR23772.1; -  
 DR EMBL: AB037667; BAA92584.1; ALT\_INTT.  
 DR EMBL: AF001697; BAA95502.1; -  
 DR EMBL: AL162080; CAB82413.1; -  
 DR PIR: T47158; T47158.  
 DR MEROPS: M12.222; -  
 DR Genew: HGNC:217; ADAMTSL1.  
 DR MIM: 605174; -  
 DR GO: GO:0008337; F:metallopeptidase activity; TAS.  
 DR GO: GO:0007229; P:integrin-mediated signaling pathway; TAS.  
 DR GO: GO:0008385; P:negative regulation of cell proliferation; TAS.  
 DR InterPro: IPR006565; ADAM\_cysteine.  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR006025; Pept\_Mn\_Zn\_BS.  
 DR InterPro: IPR001590; Peptidase\_M12B.  
 DR InterPro: IPR002870; Peptidase\_M12B\_N.  
 DR InterPro: IPR000884; TSP1.  
 DR InterPro: IPR000885; TSP1.  
 DR Pfam: PF01567; Pept\_M12B\_Tropep; 1.  
 DR Pfam: PF01421; RepPolysin; 1.  
 DR Pfam: PF00090; TSP\_1; 3.  
 DR PRINTS: PRO1705; TSP1REPEAT.  
 DR SMART: SM00608; ACR; 1.  
 DR SMART: SM00209; TSP1; 3.  
 DR PROSITE: PS50215; ADAM\_MEROPS; 1.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.

DR PROSITE: PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE: PS50092; TSP1; 3.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 KW Hydroxylase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KW Repeat: Extracellular matrix; Heparin-binding.  
 FT SIGNAL 1 49  
 FT PROPEP 50 252  
 FT CHAIN 253 967  
 FT DOMAIN 253 475  
 FT DOMAIN 476 559  
 FT DOMAIN 559 614  
 FT DOMAIN 617 724  
 FT DOMAIN 725 849  
 FT DOMAIN 854 905  
 FT DOMAIN 908 967  
 FT SITE 843 846  
 FT METAL 401 198  
 FT ACT SITE 401 401  
 FT METAL 402 402  
 FT METAL 405 405  
 FT METAL 411 411  
 FT CARBOHYD 547 547  
 FT CARBOHYD 720 720  
 FT CARBOHYD 764 764  
 FT CONFLICT 227 227  
 FT CONFLICT 468 468  
 FT CONFLICT 561 561  
 SQ SEQUENCE 967 AA; 105383 MW; C18939324741ED1 CRC64;  
 S -> H (IN REF. 1).  
 O -> N (IN REF. 1).

Query Match 48.0%; Score 2326.5; DB 1; Length 967;  
 Best Local Similarity 49.7%; Pred. No. 6; Ee-117;  
 Matches 452; Conservativity 146; Mismatches 249; Indels 63; Gaps 22;

QY 11 LPFLLLLLLLPARAPAPAPAGAGSELVVP--TRLPGSGELALHLSAFKGVLR 68  
 DB 33 VPTLLLLAALLAAS-DALGRPSE--EDELVPLELRAAGH-GTTLRLHAPQQLDLE 88  
 QY LAPDSFLAPFKIERLGSGGRATG-----ERGLRCFPGTVNGEPESLAASLCRGL 123  
 DB 89 LRPSFLAPFTLQNV---GRKSGSEPLPENTLANCFSGTVNGDPSAALSLCEGV 145  
 QY 124 SGSLLLGEERTTLPQAGGSLAPPHLQWGPAGAPLP-----RGSEMEV---ET 172  
 DB 146 RGAFYLLGEAYFIPLPAAS-----ERLATAPGEKPPAPLQHLRRNRQGVGTGCV 200  
 QY 173 GEGORERGHQDSE 222  
 DB 201 VDDEPRPRGKAFETDEDEBGTGEDEBGPQWSPQDPAQGVGPSTGIRKKRPFVSHRYV 260  
 QY 223 ETLVADASMAFYGADLQNHILTLMSVAARIYHPSIKNSINLMVYKVLIVDEKMGPE 282  
 DB 261 ETLVADQSMALFEFGSGIKHILTLFSVAARLYGHPISIRNSVSLVAVKILIVIHQSGPE 320  
 QY 283 VSDNGGLTLRFNCMORBFNPOPSDRHPHEHYDTALLTRONFCQEGGLCDTLGVADIGTIC 342  
 DB 321 VTSNALTLRFNCMORBFNPOPSDRHPHEHYDTALLTRONFCQEGGLCDTLGVADIGTIC 379  
 QY 343 DPNKSCSYIEDGLQAHTLAHELGHVLSMPEHDSKPCSTRIFGPMGKHVMAPIFVHLNQ 402  
 DB 380 DPNKSCSYIEDGLQAHTLAHELGHVLSMPEHDSKPCSTRIFGPMGKHVMAPIFVHLNQ 439  
 QY 403 TLPSPCSAWYLLTLGCGHDCCLDAPGALPLPTGPGMALYQDQCGROIFGPDFR 462  
 DB 440 SQMSPGSAWYLLTLGCGHDCCLDAPGALPLPTGPGMALYQDQCGROIFGPDFR 497  
 QY 463 HCPNTSAQVCAQMLC-HTDGAEPILCHTKNGSLFWADGTPCGPHLCSGSLPEEVEVER 521  
 DB 498 HCPD--AASCTSLIMCTGSGVLVQTKH--FPWADGTSGBGKWCINCKVAKTRKH 553  
 QY 522 PKPVYDGMAMWGMGBESRRCGGGVQFSHRECDPEPONGRCYCLGRAYQSGCHTEEC 581  
 DB 554 FDTFPHGWMGMGBESRRCGGGVQFSHRECDPEPONGRCYCLGRAYQSGCHTEEC 613





```

FT DOMAIN 213 436 METALLOPROTEINASE.
FT DOMAIN 437 519 DISINTEGRIN-LIKE.
FT DOMAIN 520 575 TSP TYPE-1.
FT DOMAIN 577 685 CYS-RICH.
FT DOMAIN 686 837 SPACER.
FT DOMAIN 837 252 POLY-ALA.
FT SITE 194 194 CYSTEINE SWITCH (POTENTIAL).
FT METAL 361 361 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 362 362 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 365 365 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 371 371 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 68 68 N-LINKED (GLUCAC. .) (POTENTIAL).
FT CONFLICT 77 77 A -> T (IN REF. 1).
FT CONFLICT 626 626 A -> Q (IN REF. 3).
FT CONFLICT 682 682 G -> R (IN REF. 3).
SQ SEQUENCE 837 AA; 90224 MW; 5DF9C9AC137DF41F CRC64;

Query Match 38.0%; Score 1842.5; DB 1; Length 837;
Best Local Similarity 45.0%; Pred. No. 7,7e-107;
Matches 375; Conservative 130; Mismatches 244; Indels 85; Gaps 24;

```

```

QY 758 ILVKGCTI-LKXSGSIAITRLQSRPLPEPTVGLLVPGEVPPKXKVFEP 810
DB 751 VLVPAVSLRYSGATPAASETSGHGFLAQPLTLVLV-VAGNPQTRLRYSFVP 803

RESULT 8
ATSS_MOUSE
ID ATSS_MOUSE STANDARD; PRT; 930 AA.
AC Q9R001.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ADAMTS-5 precursor (BC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS5) (Aggrecanase-2)
DE (ADMP-2) (Implantin).
GN ADAMTS5.
ON Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99395124; PubMed=10464288;
RA Hukakinen T.U., Hirohata S., Seidlin M.F., Appe S.S.;
RT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of
RT zinc metalloproteinases."
RL J. Biol. Chem. 274:25555-25563 (1999).
CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
CC involved in its turnover. May play an important role in the
CC destruction of aggrecan in arthritic diseases. May play a role in
CC proteolytic processing mostly during the peri-implantation period.
CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-Ala-393
CC site.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -1- DEVELOPMENTAL STAGE: Expressed specifically in the peri-
CC implantation period in embryo and trophoblast and at low or
CC undetectable level thereafter.
CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -1- SIMILARITY: Belongs to peptidase family M12B.
CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
CC -1- SIMILARITY: Contains 2 TSP type-1 domains.
CC -1- SIMILARITY: Contains 2 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
CC EMBL: AF140673; AAD56356.1; -.
CC MEROB: M12.225; -.
CC MGD: MG1:1346321; Adamts5.
CC InterPro: IPR001762; Disintegrin.
CC InterPro: IPR006025; Pept_M12B.
CC InterPro: IPR001590; Peptidase_M12B.
CC InterPro: IPR002870; Peptidase_M12B_N.
CC InterPro: IPR000884; TSP1.
CC InterPro: IPR008085; TSP1.
CC Pfam: PF01562; Pp_M12B_propep. 1.
CC Pfam: PF01421; Repolysin; 1.
CC Pfam: PF00090; TSP1.2.
CC PRINTS: PR01705; TSP1RBPAT.
CC SMART: SM00209; TSP1.2.
CC PROSITE: PS00215; ADAM_MERO. 1.
CC PROSITE: PS00392; TSP1. 2.
CC PROSITE: PS00142; ZINC_PROTEASE. 1.
CC PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.

```

KM Hydrolase; Metalloprotease; Zinc; signal; Glycoprotein; Zymogen;  
 KM Repeat; Extracellular matrix.  
 FT SIGNAL 1 21  
 FT PROPEP 22  
 FT CHAIN 261  
 FT DOMAIN 262  
 FT DOMAIN 262  
 FT DOMAIN 484  
 FT DOMAIN 485  
 FT DOMAIN 567  
 FT DOMAIN 624  
 FT DOMAIN 732  
 FT DOMAIN 874  
 FT DOMAIN 875  
 FT DOMAIN 929  
 FT SITE 41  
 FT SITE 257  
 FT METAL 209  
 FT ACT SITE 410  
 FT METAL 411  
 FT METAL 414  
 FT METAL 414  
 FT METAL 420  
 FT CARBOHYD 498  
 FT CARBOHYD 498  
 FT CARBOHYD 728  
 FT CARBOHYD 802  
 FT CARBOHYD 807  
 SQ SEQUENCE 930 AA; 101780 MM; 84DE84B26170D4DC CRC64;  
 Query Match 36.5%; Score 1769.5; DB 1; Length 930;  
 Best Local Similarity 40.3%; Pred. No. 2.9e-102;  
 Matches 381; Conservative 124; Mismatches 355; Indels 85; Gaps 19;  
 10 WLPFLILLILLL---LPLPARGA---RPAAGQASLVLV-----44  
 5 WASLILLILLILLLASCLSLAASPAADKTRQQAALAAEPDQGEETREHQLQ 64  
 45 -----TRLPASAGELALHLASFGKGVFLIAPDPSFLAEFKIERLGSGR 90  
 65 LAGGRSGGLVNIIDQLVSGGKGVLYVAGGRRLDLERDT-VGAAGSYVAGGLS 123  
 91 ATGGERGLRGCFSGTNGEPELSLAVALICRLSGSFLIDGEBFTIQGAGGSLAQR 150  
 124 ASSGRHG-HCFRGTVDSPRSLVFDLCGLDFFFAKARLYLKLPL-LRGSVAEYER 180  
 151 LORWGPAGAR-----PLPRGMEVEVEGEQORERGDHQDESEESQ-----192  
 181 I---YDGGSRILHYNNRBSFEALPPRASCTPASPBPQESPVHSRRRLAPQL 238  
 193 -EEBAGASPEPPPGATSRTPFVSEAPVETLLVADASMAAFYADLQNHILLMSVA 251  
 233 LDHSAFSPSGNAGPOTWRRRRRSISRAROVELLIVADSSMARMGRLQHYLLTLASIA 298  
 252 ARIYKHPSTIKNSINIMVYVLLIVEDEKMGPEVSDNGSLTLNFCMORRFPQSPRPH 311  
 299 NLYSHASIEHRIKLAIVKVVLTDKDTSLEVSXKAAATTLKFCMKQGHQNLQDDBEH 358  
 312 YDAILLRONFCGEGELCDTLGVADIGTICDPNKSQSVIIEDEGLQAHTLAHELGHVLS 371  
 359 YDAALIFREDICGHHN-CDTLGMADVIGTICSPESCAVIEDDGHAAFAVHAHEIGHILG 417  
 372 MEHDSKPEPTRLFGPMGKHVMAPLFVHNOTLPPSPASAWLTLEILDGGDCLDAPG 431  
 418 LSHDSKCEENFGTTEDEKRLMSLITSLDASKPISKTSAITIEFLDGGNCLDLPR 477  
 432 AALPLPTGLPGMALYOLDQGCROIFGPDRHCPTNSADQVCAOLMCH-TDGAPLCHTK 490  
 478 KQILPEELPQO--TYDATQCNTLTFGEYSVCP--GMDVCARLMCAVVAQGVVCLTK 532  
 491 NSLSMAOGTGPCGPHLQSESGCLPEBEVERKPVVDGMAFWMGWEGSCSTTCGGVQFS 550  
 533 --KLPAVBSTPCGKGRVCLQCKVCKTKKYYSTSSHANWGSWGWGCGSSCGGQVFA 590  
 551 HRECKDPENONGRAYCLGRAYQSCHEBPCPDGSPREDOCEKYNAYNTMDG--NL 608  
 591 YAHGNPAPRNSGRCTGKRAIYRSCSVTPCPBNKXSRFHQCEAKNGVQ--SDAKGYTFF 649  
 609 LOWPKYAGVSPRDRCKLFCRAKGRSBEKVFVAKYIDGTLGPELTALCTVAGQCVKAGCD 668

Db 650 VEWPKTAGVLPADVCLTRPRAGTGYVVFSPYVDGTGCRYSNVCRGRCVTCGD 709  
 Qy 669 HVVDSPRKLDKPCGVCGKNGNSCKRVSGSLTPTYGVNDIYTPAGATNIDVKRSHBGVQ 728  
 Db 710 GIISKIQYKCGKCGCGDNNSSCTKIKTGFNKKSQGYDVIRIEGATHIVRQPKADQT 769  
 Qy 729 NDGVYALKTKADQGYLLNGMLAISALEQDLVYGTILKSGSLATLERL--QSFRPLPEP 786  
 Db 770 RFPAYLALKKKTBEYLNGKRMISTSETIIDINGTVWNTSGWHRDPLHGMGSATKEI 829  
 Qy 787 LTVQLTVPGEVPEPPK---VKYTFEVPNDVDFSMQSSKERATTNIIQPLHAQVIGDMS 843  
 Db 830 LIQILATD---PTALGVRYSFVPKTKTKVNSVISHGSKVGHSTQLQWVGPWL 885  
 Qy 844 EGSSTCAGMQRTVEGRDPSGQASATCNKALPEADKPCESQLC 888  
 Db 886 ACSRTCDTGMHRTVOCODGNRLKAGCLLSQPSAFKQCLTKKC 930  
 RESULT 9  
 ACSS\_HUMAN  
 ID ACSS\_HUMAN STANDARD; PRT; 930 AA.  
 AC Q9UNAO; Q9UNP2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE ADAMTS-5 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 5) (ADAM-TS 5) (ADAM-TS5) (Aggrecanase-2) (ADMP-2) (ADAM-TS 11).  
 GN ADAMTS5 OR ADMP2 OR ADAMTS11.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=9367476; PubMed=10438522;  
 RA Abbaasade I., Liu R.-O., Yang F., Rosenfeld S.A., Ross O.H., Link J.R., Ellis D.M., Tortorella M.D., Pratta M.A., Hollis J.M., Wynn R., Duke J.L., George H.J., Hillman M.C. Jr., Murphy K., Mswall B.H., Copeland R.A., Decicco C.P., Bruckner R., Nagase H., Ito Y., Newton R.C., Magolda R.V., Trzaskos J.M., Hollis G.F., Arner E.C., Burn T.C.;  
 RA "Cloning and characterization of ADAMTS11, an aggrecanase from the ADAMTS family.";  
 RL J. Biol. Chem. 274:23443-23450 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20289799; PubMed=10830953;  
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y., Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W., Rosenblatt A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E., Minoshima S., Shimizu N., Nordstrem G., Hornischer K., Brandt P., Scharte M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecher H., Ramser J., Beck A., Klages S., Hennig S., Rlesseimann L., Dagand E., Mehmeier S., Borzym K., Gardiner K., Nizetic D., Francis F., Lebrun H., Reinhardt R., Yaspo M.-L.;  
 RA "The DNA sequence of human chromosome 21.";  
 RL Nature 405:311-319 (2000).  
 RN [3]  
 RP SEQUENCE OF 413-930 FROM N.A.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=9395124; PubMed=10464288;  
 RA Hurskainen T.L., Hirohata S., Seldin M.F., Apte S.S.;  
 RT "ADAM-TS, ADAM-TS6, and ADAM-TS7, novel members of a new family of zinc metalloproteases.";  
 RL J. Biol. Chem. 274:25555-25563 (1999).

CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
 CC involved in its turnover. May play an important role in the  
 CC destruction of aggrecan in arthritic diseases. May play a role in  
 CC proteolytic processing mostly during the peri-implantation period.  
 CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 392-Glu-Ala-393  
 CC site.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (By similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed at low level, primarily in placenta  
 CC but also in other tissues, such as heart and brain, and also  
 CC cervix, uterus, bladder, esophagus, rib cartilage,  
 CC chondroblastoma, fibrous tissue and joint capsule from an  
 CC arthritic patient.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix.  
 CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 2 TSP type-1 domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL; AF142039; AAD9577.1; -  
 DR EMBL; AP001688; BAA95504.1; -  
 DR EMBL; AP001697; BAA95503.1; -  
 DR EMBL; AF141293; AAF02493.1; -  
 DR HSPB; O9PM35; 1BD.  
 DR MEROPS; M12.225; -  
 DR Genew; HGNC:421; ADAMTS5.  
 DR MIM; 605007; -  
 DR GO; GO:0005578; C:extracellular matrix; TAS.  
 DR GO; GO:0005178; F:integrin binding; TAS.  
 DR GO; GO:0008237; F:metallopeptidase activity; TAS.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006025; Pept\_M12B.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP 1.  
 DR Pfam; PF01562; Pept\_M12B\_propep; 1.  
 DR Pfam; PF01421; Reptolysin\_1.  
 DR Pfam; PF00090; TSP1\_2.  
 DR PRINTS; PRO1705; TSP1REPEAT.  
 DR SMART; SMO0209; TSP1\_2.  
 DR PROSITE; PS50215; ADAM\_MEROPS; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE NEG.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE NEG.  
 DR PROSITE; PS50092; TSP1\_2.  
 DR PROSITE; PS00143; ZINC\_PROTEASE; 1.  
 DR HydroLase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KM Repeat; Extracellular matrix.  
 FT SIGNAL 1 16  
 FT PROPEP 17 261  
 FT CHAIN 262 930  
 FT DOMAIN 262 484  
 FT DOMAIN 485 566  
 FT DOMAIN 567 622  
 FT DOMAIN 624 731  
 FT DOMAIN 732 874  
 FT DOMAIN 875 929  
 FT DOMAIN 930 941  
 FT DOMAIN 942 957  
 FT SITE 257 261  
 FT SITE 209 209  
 FT METAL 410 410  
 ZINC (CATALYTIC) (BY SIMILARITY).

FT ACT SITE 411 411 BY SIMILARITY  
 FT METAL 414 414 ZINC (CATALYTIC) (BY SIMILARITY)  
 FT METAL 420 420 ZINC (CATALYTIC) (BY SIMILARITY)  
 FT CARBOHYD 498 498 N-LINKED (GLCNAC...) (POTENTIAL)  
 FT CARBOHYD 728 728 N-LINKED (GLCNAC...) (POTENTIAL)  
 FT CARBOHYD 802 802 N-LINKED (GLCNAC...) (POTENTIAL)  
 FT CARBOHYD 807 807 N-LINKED (GLCNAC...) (POTENTIAL)  
 FT CONFLICT 138 138 A -> G (IN REF. 2)  
 FT CONFLICT 614 614 R -> H (IN REF. 3)  
 FT CONFLICT 692 692 P -> L (IN REF. 2)  
 SQ SEQUENCE 930 AA; 101715 MW; B64281502F28193B CRC64;  
 Query Match 35.7%; Score 1734.5; DB 1; Length 930;  
 Best Local Similarity 39.8%; Pred. No. 4.3e-100;  
 Matches 378; Conservative 122; Mismatches 348; Indels 101; Gaps 20;  
 15 LLLLLLLPLARGPAPPA---GGC---ASLVVPRRLG----- 49  
 8 LLLCAFLPLAAVGPATPADQXAGPPTAAAAQPRRQGEVQERAPPPHPLAQR 67  
 50 -----SAGELALHSARFGKGVFLRADPDSFLAPDFKIERLGSGRATGG 94  
 68 RRSKGVNIDQLYSGGKGVLYVYAGRRLLDLERDGSV-----GIAGVPVAGG 119  
 95 E---RGIRGCFRGTVNGEPESLAVALSLCRLSGSLDGEFTIQOGAGSLAOPHR 150  
 120 TSAPFRHSHFHYRGTVASPRSLAVFDLCGLDGFPAVKARATLSP-LRGWAEBEK 178  
 131 LQRMGPAGARPLPGPEVEVETGSGRQE---RGDQDEBEBSQEEAEAGSEPP--- 203  
 179 GRVYDGSARLL-----HYTRRGFSFEALPPRASCTEPASTEAEHHAHNPGR 232  
 204 -----PLGATG-----RTKRFVSEARFVETLLVADASAAAFYADLQNHIL 245  
 223 ALASQLDQASLSPAGSGGQPTWRRRRRSISRQVELLVADASAAAFYGRGLQHYLL 292  
 246 TLMSVAALIVGSPILKYSINLMVYKVLVDEKKKGPVSDGGLTLNFCMORRFPQS 305  
 293 TLASTANLSYHASTENHRLAAVAVVVLGDKSLVSKAAATTLNFCMORHQLG 352  
 306 DRHPEHYDTALLTRNFCQGEGLCDTLGVADIGTICDPNKSQVIBDEGLQAHTLAHE 365  
 353 DDHEHYDAALILFRIDLCGHS-CDTLGMADVGTICSPERSCAVIDDDGHAFFVAHE 411  
 366 LGHVTSMFHDSKCTRLFGPMGKHVMAPEFVHLNQTLPSPSPSANYLTLELDCGHDC 425  
 412 IGHLLGSHDSKCEETFEISTEDKRLSSILTSIDSKPWSKTSATTEFLDDGHNC 471  
 426 LLDAFGAALPLPTGLGFMALYOLDQOCROIFGDFPHCPNTSAQDYCAOLMCH-TDGA 484  
 472 LLDEPRQLLPBELPQ--TYDANQCNLTFGEYSVCP---GMDVCARLMCAVVRQG 526  
 485 PLCHTKNSLWMAOGTPGPHLCSBGSCLPEEVEVERPKVYVDGMAPMGWGCSPTCG 544  
 527 WCLTK-KLAAVGTGPGKGRICLQGVCKTKKRYSTISSHGNWSGWSGQCSHSCG 584  
 545 GGVQFSHRECKDPEPONGRYCLGRRAKYQSCHTBEPDPDGSFREOCCEKXNYNTDM 604  
 585 GGVQAYAHNCPNAPRNNGRYCTGKRALYRCSLMPCEPNKSRHCEAKNGYQ-SDA 643  
 605 DG--NLQWVRKXVGSFRDCKLFCARAGSBEKVPEAKYIDGTLCGPETLALCVNGQC 662  
 644 KGVKTFVEMVVKIGVLPADVCKLTCAKRGYGVVSPKTYDTECPYNSVCAVNGKC 703  
 663 VYAGCDHVVDSPRLDKVCGGKSGSKYSGSLTFPNYGYNDIVITIPAGATNIDVYOR 722  
 704 VRTGDDGITSGLQYDKGVCGGDSSTCKLVGFNNKSKKGYTVLRPEATHIKRQF 763  
 723 SHPGVQDGNVYALKTADGQYLNGNTAISALBODILVKGITLKYSGSLATLEL--QSF 780  
 764 KAKODTRFTAVYALAKKNGEYLINGKYMISTEYIILINGVNNYSGSHRDEFLHNGY 823  
 781 RPLPEPLTVQLL-TVPGSVFPKXVYTFEVENDVDFMSQSKERATNIIQPLHAQVVL 839





Db 142 TAVFSLGGIMGTGFKANDGEYFLEPYLRADGSAHDDHKKPHLYR-----QELKRNSE 195  
 Qy 166 --PEWEVETGEGORORGDHDESEESQ-----EEEA---EGASEPPPLGATSRKTF 215  
 Db 196 ARBKHCCESENMETALPSSQSSRTTGDVDEEAVALRLEBSRQ-----LHSHKRF 250  
 Qy 216 VSEARFVETLLVADASMAAFYGADLONHILLMSVARYKPSIKNSINLMVAVYLYE 275  
 Db 251 LSYRYRYEVWVTLDAKVAHHGQNLQHYVTLMSIYAAIYKSSIGNLNIYIVKLVVH 310  
 Qy 276 DEKMGPEVSDNGGLTLRNFCMQRFRNPSDRPEHYDTAILLTRNFCGSEGLCDTLGV 335  
 Db 311 SEGBGVPTSNATTLNPLNMQSQSNVDPDAPRSHHDVALITRBDIGAKKEDTGL 370  
 Qy 336 ADIGTICDPKSCSVIIEDEGLQAAPLAEHLGYLSNPHDSCPCRTLRFGMGKHHVAP 395  
 Db 371 AELGTICDPKSCSVISEENGLSAFTLAHELGHVFNPHDSCPCKEA-GIKQHYVAVAP 429  
 Qy 396 LFNHNLTLNPSPCSAWYLTLLDGGHGDCLLDAP-GAALPLPTGPGSMALYOLDQOR 454  
 Db 430 TLNHTSPWTWSACSQKHTEFLDTGHGECILDKPKNRITDLSPLQPG--SVYDNKQCE 487  
 Qy 455 QIFGPDPRHCPNTSADVCAQLWC-HTDGAELPCHTKNGSLPWADGTCGPGHLCSESGC 513  
 Db 488 LMFPGSQVCPYLKH---CRLWCTSAEGVHKGCRTQH--MPLADGTSQGPWGHCHRGIC 542  
 Qy 514 LPEEVEVREKPVVDGGAHPMGWGECSRTGGGVQFSHRCNDREPONGRVCIGSRAX 573  
 Db 543 VTRDMETRP--VDGEWGPWGPYSSCSRTGGGISTARLCDPEPRNGRVCVGRMKF 599  
 Qy 574 QSCHTCEPPDGKSFREQCEKYNANYITMDG--NLLQWVPKAGVSPDRCKLFCRAR 631  
 Db 600 RSCNTDCEPKGKDFREKQSDPDGKH-F-DINGLPVAVRMLPKYSGIAYVDRCKLYCRA 658  
 Qy 632 GRSEPKVFEAKVIDTLCGPETLAIYRGQCVAKGCDHVDSPRKLDKGVCGGKNSCR 691  
 Db 659 GTTSFYOLKDRVADGTPCGTEINDICVQGLCRQAGCHVANSKAKDKCVCGGDSISQ 718  
 Qy 692 KVSGLSLPTNYGVNDIYTPAGATNIVKORSHGVQNDNYALAKTADQYLLNGLAI 751  
 Db 719 TLAQVNSAHYGVNVVXKIPGATNITIILOHSYSGREDNYIALSTQGFLLNGLFV 778  
 Qy 752 SAIBODILVKTILKYSGSIATTLERLQSFRLPRLPVOLLTPGVEVFPKXKTFVP 810  
 Db 779 SMAKEINIQAVVEYSGNSISIRINSTRLAEALVQLCY-GMLYNPDVYSPNIP 837  
 Qy 811 -----NVDPSMOSKERATT-----NIIQPL----- 833  
 Db 838 EERSNLPSWDYGMQDCTTKKQGLHRRKIACTRKSDHAVVSDHNGHLMPLFTEKCN 897  
 Qy 834 --HAQWYLGDMSECSSTCGAGMQRRTVEGRDPS--GQA-----SATCNALKPEKAPC 863  
 Db 898 MDCLRAHIIIGKSSCGGCGGYTLVDYHCKYGVHKGQAVPVQDYGQDLKPRSPREPC 957  
 Qy 884 ESQICPL 890  
 Db 958 HGS-CVL 963  
 RESULT 11  
 AT99 HUMAN  
 ID AT99 HUMAN STANDARD; PRT; 1935 AA.  
 AC Q9P2N4; Q9NR29;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE ADAMTS-9 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 with thrombospondin motifs 9) (ADAM-TS 9) (ADAM-TS9).  
 GN ADAMTS9 OR KIAA1312.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;

RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 3).  
 RC TISSUE=Fetal;  
 RX MEDLINE=20396138; PubMed=10936055;  
 RA Clark M.E., Kellner G.S., Turdeville L.A., Boyer A., Arden K.A.,  
 RA Maki R.A.;  
 RT "ADAMTS 9", a novel member of the ADAM-TS/Metalloprotein gene  
 RT family.";  
 RL Genomics 67:343-350(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND FUNCTION.  
 RX MEDLINE=2251925; PubMed=12514189;  
 RA Somerville R.P., Longpre J.-M., Jungers K.A., Engle J.M., Ross M.,  
 RA Evanko S., Wright T.N., Leduc R., Apte S.S.;  
 RT "Characterization of ADAMTS-9 and ADAMTS-20 as a distinct ADAMTS  
 RT subfamily related to Caenorhabditis elegans GON-1.";  
 RL J. Biol. Chem. 278:9503-9513(2003).  
 RN [3]  
 RP SEQUENCE OF 159-1935 FROM N.A. (ISOFORM 2).  
 RC TISSUE=Brain;  
 RX MEDLINE=20181126; PubMed=10718198;  
 RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XVI.  
 RT The complete sequences of 150 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 7:65-73(2000).  
 CC -1- FUNCTION: Cleaves the large aggregating proteoglycans, aggrecan  
 CC and versican.  
 CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1838-Glu-Ala-1839  
 CC site and versican at the 1428-Glu-Ala-1429 site.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (By similarity).  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing: Named isoforms=3;  
 CC Name=1; Synonyms=ADAMTS-98;  
 CC IsoId=Q9P2N4-3; Sequence=Displayed;  
 CC Name=2; Synonyms=Long;  
 CC IsoId=Q9P2N4-1; Sequence=VSP\_007548, VSP\_007549;  
 CC Note=May result from the retention of an intron in the cDNA  
 CC leading to a premature stop codon;  
 CC Name=3; Synonyms=Short;  
 CC IsoId=Q9P2N4-2; Sequence=VSP\_005499, VSP\_005500;  
 CC -1- TISSUE SPECIFICITY: Highly expressed in all fetal tissues.  
 CC Expressed in a number of adult tissues with highest expression in  
 CC heart, placenta and skeletal muscle.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix (By  
 CC similarity).  
 CC -1- DOMAIN: The ancillary domains, including the TSP domain, are  
 CC required for specific extracellular localization and for its  
 CC versicanase and aggrecanase activities.  
 CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M2B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 1 GON domain.  
 CC -1- SIMILARITY: Contains 15 TSP type-1 domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; AF261918; AAF89106.1; -  
 CC EMBL; AF488803; AA015765.1; -  
 CC EMBL; AB037733; BAA92550.1; -  
 CC HSSP; P15167; IATL.  
 CC MEROPS; M12.021; -  
 CC Genew; HGNC:13202; ADAMTS9.





ID AT54 RAT STANDARD; PRT; 630 AA.  
 AC Q9ESF7; Q9ESF6; Q9ESF8;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE ADAMTS-4 precursor (EC 3.4.24.82) (A disintegrin and metalloproteinase with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1) (Fragment).  
 DE ADAMTS4.  
 GN Adamts4.  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Mistar; TISSUE=Brain;  
 RX MEDLINE=20415831; PubMed=10961658;  
 RA Satoh K., Suzuki N., Yokota H.;  
 RT "ADAMTS-4 (a disintegrin and metalloproteinase with thrombospondin motifs) is transcriptionally induced in beta-amyloid treated rat astrocytes.";  
 RT Neurosci. Lett. 289:177-180 (2000).  
 RL  
 CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be involved in its turnover. May play an important role in the destruction of aggrecan in arthritic diseases. Cleaves aggrecan at the 392-Glu-Ala-393 site.  
 CC -1- CATALYTIC ACTIVITY: Glutamy1 endopeptidase; bonds cleaved include 370-Thr-Glu-Gly-Ala-Arg-Ser-377 in the interglobular domain of mammalian aggrecan.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).  
 CC -1- TISSUE SPECIFICITY: Only expressed in brain.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.  
 CC -1- PFM: The precursor is cleaved by a furin endopeptidase.  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 1 TSP type-1 domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 DR EMBL, AB042272; BAB16474.1; -;  
 DR EMBL, AB042271; BAB16473.1; -;  
 DR EMBL, AB042273; BAB16475.1; -;  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR006025; Pept\_Mn\_Bs.  
 DR InterPro: IPR001590; Peptidase\_M12B.  
 DR InterPro: IPR000884; TSP1.  
 DR InterPro: IPR008085; TSP1.  
 DR Pfam: PF01421; Reprolysin; 1.  
 DR Pfam: PF00090; TSP1; 1.  
 DR PRINTS: PR01705; TSP1REPEAT.  
 DR SMART: SM00209; TSP1; 1.  
 DR PROSITE: PS00215; ADAM\_MEPRO; 1. FALSE NEG.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE NEG.  
 DR PROSITE: PS00214; DISINTEGRIN\_2; FALSE NEG.  
 DR PROSITE: PS00092; TSP1; 1.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 KM Hydroxylase; Metalloprotease; Zinc; Glycoprotein; Zymogen; Extracellular matrix.  
 KM  
 FT NON\_TER 1  
 FT PROPEP 5  
 FT CHAIN 6 630 BY SIMILARITY.  
 FT DOMAIN 6 232 ADAMTS-4.  
 FT DOMAIN 233 303 METALLOPROTEASE.  
 FT DOMAIN 313 368 DISINTEGRIN-LIKE.  
 FT TSP TYPE-1.

FT DOMAIN 368 478 CYS-RICH.  
 FT DOMAIN 479 630 SPACER.  
 FT DOMAIN 630 630 POLY-ALA.  
 FT METAL 154 154 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 155 155 BY SIMILARITY.  
 FT METAL 158 158 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 164 164 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 630 AA; 68384 MW; 63A428753167C7E7 CRC64;  
 Query Match 33.9%; Score 1644.5; DB 1; Length 630;  
 Best Local Similarity 49.9%; Pred. No. 1e-94; Indels 21; Gaps 13;  
 Matches 305; Conservative 107; Mismatches 178;  
 211 RTKRFVSEARFEVETLVADASMAAFYGDLONHITLMSVARIYKHSINSLMYVK 270  
 2 RTKRFASLSRFVETLVADDDKAAAFHAGLKHVLLTVAALAAKAFKHSIRNPVLVYTR 61  
 271 VLIIVDEKMGKPVSDNGSLTLNFCNMGRNPNPSDRPHHYDTALLTRNFCGQEGLC 330  
 62 IVLISGGEVPPQVGSAAQTLSFCTWQKGLNPPVSDPDHDTALLTRDQLG-VSTC 120  
 331 DTLGVADICTIDPKKSCSVIEDEGLQAHTLAHELGHVLSWPHDSKPCRTLFGP-MGK 389  
 121 DALGMAVGVTCDDPARSCAIVEDDLQSAFTAHELGHVFNMLHDSKPCANLNGQSSS 180  
 390 HHVAPLFLVHLNQLTPNSPCSMYTELIDGHCCLLDAPQALPLPTGLPGRALYQL 449  
 181 RHVMAVPMVAHVDPBPWSPCSARFTIDLDNGYGLCLDKPAPLHLVTPFGKD-YDA 238  
 450 DQCGRIQFDPFRHCPNTSAQVCAQWLC--HTDGAELCHTKNGSLPMADGTPCGPQHL 507  
 239 DQCGRIQFDPFRHCPNTSAQVCAQWLC--HTDGAELCHTKNGSLPMADGTPCGPQHL 293  
 508 CSBGSLPEEVEVERKPVVDGMAVPMWBCSRTGCGVQFSRHCQDPBPONGRCL 567  
 294 CMGGRCLHVDQDKDNIPDQAGGMPWGDGSRCTGCGVQFSRHCQDPBPONGRCL 353  
 568 GBRKQSGHTEBEP-PDGKSPREOCCKRYANVYTDGNNL---LOWPKAGSPRDR 623  
 354 GRTTPRSNTNKPCHGSAITREBOCAALN--HTIDLFKSPGMDVNPRTVGAIPDQ 411  
 624 CLKFCARGSRSEFKFEAKVIDGLCGPETAICVGGCVAGCDHVVDSPKLDKGVY 683  
 412 CLKTCQARALGYVYLEPRVADGTCSPSSSVYQCGICHAQCRILIGSKKKPKVCWC 471  
 684 GKGKNSCRVSGSLPPTNAGVNDITIPAGATNIDVQRSHRGVQNDGNYALAKTADGQY 743  
 472 GNGSGSCSKQSGSFKKFRYGSVDVVTIPAGRTHILVRQGGSGLS--IYLLKLADSY 529  
 744 LINGNLAIASIEODILVKGTI-LKYSGSITATLERLQSFAPLEPPTVOLLVPGSEVPPK 802  
 530 ALNGEYTLMPSTDVLPAAVSLRYSGRRAAETISGHQPLAQPLTLQVL-VAGNPQVNR 588  
 803 VKYTFVENDV 813  
 589 LRYSFVPRPV 599  
 RESULT 13  
 AT20\_HUMAN STANDARD; PRT; 1911 AA.  
 ID AT20\_HUMAN  
 AC P59510;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE ADAMTS-20 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 20) (ADAM-TS 20) (ADAM-TS20).  
 GN ADAMTS20.  
 OS Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;



```

Db 64 QKRSSEALEPMPFRTHTYFAYGQLFQNLADASEFLAAGTVEVHLGTPERGAMESDAGP 123
QY 96 RLRCGFSSGTNGEPESLAVSLCRGSGSFLGGEFTTOP-QGAGSILAQHRLORW 154
Db 124 SDRHCFYRGQVNSOEDYKAVVSLGGGLGTGKQNGEYFLEPIKADGN----- 173
QY 155 GPAGARPLPRGPEWEVETGEGORERGDHOED-----SEESQ----- 192
Db 174 -----EYE-----DHINKPHLITRODLNNSFLQTLKYCSVSSQIKENSLPHT 217
QY 193 -----EBEAGASEPPPEL---GATSTRKPFVSEARFVETLLVADASMAAFGA 238
Db 218 YENMEDLVMKERVIGHTSKNVPLKDERHRSKRRLISYPRYIIMVTAJAKVVSAGS 277
QY 239 DIONHITLMSVAARLYKHPSTKSNINIMVAVLLIVEDEKGPEDSDNGGLTLRFQNMQ 298
Db 278 NQNTYITLMSIVATYKOPSLGNLHIVVVLVIMHREBGPVINFQATTLKNFCGSMQ 337
QY 299 REFNPSPDRHPEHYDTAILLTRONFCGDEGLCDLTVADIGTICDPNKSQVIEDEGLQA 358
Db 338 QTONDLDDVHPSHDITAVLITREDICSSKEKCNITGLSYLGTICDPLQSCFINEKGLIS 397
QY 359 AITLHAEHLGVLSMHPDSDKPCRTLFGPMGKHMAAPLVHLNQTLFSPSASAMYLELL 418
Db 398 APTIAELHGTTLGVQHDNPRCKEM--KVTKXHVAPALSPFMSWMSNSCKRKYVIFPL 455
QY 419 DGGHGDCLLDAPGAL-PLPTGLPGSMALYOLDQOCROIFGDPFHCPNTSADQVCAQIM 477
Db 456 DRYGQCLDKDEBEIYNLPSELPG--SYDQNKQCELAFGGQWCPHI---NICMLM 510
QY 478 C-HTDAEPLCHTKNGSLPMADGTPGPGPHLCSBGSCLPEEVEERPKPVQSGMAPKWP 536
Db 511 CSTETKXHKGCFFQH--VEPADGTGCPMKHCRHGLCNKKEETEP---VNSEWMPMEPY 565
QY 537 GECSTFCGGGVOPFSEHCEKDPONRGRCYCLGRAYQSCHTBECPPDGSFREOCCEKY 596
Db 566 SSSSRKTCGGGIESATRCNRPBRNGNVCVRHMKFRSCNDCSPKQDPRHQCQDF 625
QY 597 NAYNTYDMDG--NLLQWPKYAGVSPRDCCKFCFARAGSEKRVBEAKYIDTLCGPETL 654
Db 626 NG-KHLDIGSIPSNVMLPRYSGLGKDRCKLYCQVAGNIFYLLKDWEDGTGCTETH 684
QY 655 AICVRGCVYAGCDHYVDSPRKLDKCGVCGKGNCSKRYSGSLPTNYGYNDIVITPAGA 714
Db 685 DTCVQGCMAAGCDHYLNSAKIDKCGVCGGDNSSCKITTVGFNSHGYVNVVXIPAGA 744
QY 715 TNIDVQSRHPGVQNDGNIALTKTADQYLNGNLAISAIEDILVKG--TILKTSGLIA 772
Db 745 TNVDIQSYSG--QPDSDYALASDAENFLFNGLSTSKKEINQGRIVIEYSGSN 803
QY 773 TIERLQSFRLPEPLVQVLLTVGGEVFPKVKYTFVP----- 810
Db 804 AVERINSTNQEKEILLIQVLCV-GNLYNPVHYSFNIPLEKSDMTWTPYGMWECTGM 862
QY 811 -----NDVDFSMQSKRE----- 822
Db 863 CQGLQRNITCIHKSDHSVVDKEDHLPLPSFVTSQNTDCELRMLAKTSCGSSQCG 922
QY 823 ----- 822
Db 923 QGYRTLDHCKMYSIHGGQVVDHYGDQKLPFQELCHGNCVTRHMYSEWQCSRS 982
QY 823 -----RATTNIIOPLLHA-----QWVLGWSSECSSTGAGW 853
Db 983 CGGGERESREYCMNNFHLADNECBLSVTRBENGNEFCSPMAEWSSECLVTCGKGT 1042
QY 854 QRTVECR-DPSGQASATCKALKPEPAKCESQLC 888
Db 1043 KORQVWQMLVDHLSDFCNSSSTRPESLSPCELHTC 1078

```

```

ID AT15_MOUSE STANDARD; PRT; 562 AA.
AC PS93B4;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-15 (EC 3.4.24.-) (A disintegrin and metalloproteinase with
DE thrombospondin motifs 15) (ADAM-TS 15) (ADAM-TS15) (Fragment).
GN ADAMTS15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RX MEDLINE=2388257; PubMed=12477932;
RA Kraussner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Wax S.L., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.J.,
RA Bata S.S., Logucliano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Paley J., Helton E., Ketterman M., Madan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywinski M.I., Skalek U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix (By
CC similarity).
CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -1- SIMILARITY: Belongs to peptidase family M12B.
CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
CC -1- SIMILARITY: Contains 3 TSP type-1 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; BC043308; AA043308.1; -
CC MGD; MGI:2449569; Adamts15.
DR InterPro; IPR006586; ADAM cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR001819; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M2n_BS.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR Pfam; PF00090; Tsp_1; 3.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS0215; ADAM_MPRO; PARTIAL.
DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.

```



Fri May 7 12:18:03 2004

us-09-989-687-4\_1.rsp

Page 20

Best Local Similarity 31.7%; Pred. No. 6.4e-67;  
Matches 329; Conservative 137; Mismatches 369; Indels 202; Gaps 42;

```
QY 3 PAPAAPWLPPLLLLLLLPLARG---APAPAAGGQASLVLPRLPGSAGELALHL 58
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 PPSRSP--APLLRPLLLLCALAPGAPGAPGRATEGRALDIVHVRVADAGSFLSYEL 62
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 59 -----SAF-----GKGFVLRAPDPSFLAPEFKIE--RLGSGGRATG 94
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 63 WPRALARKDVSRDAPAFYELQYRGRELRFNLNLANHLLAPGFVSETRRGGLGRA--- 119
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 95 ERGLR---GCFPGTNGEPE--SLAASLCRGLSGSFLLDGEFTTQPGAGGSL-- 145
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 120 --HRAHTPACHLGEVQ--DELEGGLAALSACDGLKGVFLSNEDYFLEPLDSAPAPRG 176
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 146 -AOPRLQRMGPAGARPLPRGPEWEVEVTEGQROERGDHODESESESEAEAGSEPP 204
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 177 HAQPHVYVK-----RQAPERLADGDSASASTGCVYPELSRRERWQQRQWR 227
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 205 PLGATSRTKRF---VSEARFVETLLVADASMAAFYG-ADLQNHILLMSVAAATYKPS 259
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 228 P-----RLRLHQREVSSEKKEVETLVVADAKWEYHGOQPVESEYVLTIMMVAGLPHDS 282
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 260 INGSINLWVVKLVVEDEKMPPEVSDNGLLTRNFCWQBRFQPSDRPHEHYDTAILLT 319
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 283 IGNFHTITVRLVLEDEEDLKITTHADNTLKSFCMKQKSIKMGDAPRLHDTAILLT 342
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 320 RQNF--GQEGLCDTLGVADITICDPKSGSVIEDEGLQAHTLAHELHLSMPHDS- 377
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 343 RKDLCAANRPCELTGLSHVAGMCPHRSINEDTGLPLFTVAHGLSHSGICQHDGSG 402
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 378 KPCTRLFGPMGKH-HVNAPLFVHLNQLTPMSPCSAMLTLLDGHGDCILDAFPA-ALP 435
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 403 NDCE---FVGRPRFIMSPQLLYDAAPLTWRCRQYITRFLDRGMLCLDDBPACDILD 458
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 436 LPTGLPGMALYQLDQCCRQIFPPDFRHCNPTSAQDVCAOLKCHTDGAEPLCHTKNGSLP 495
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 459 PPSVPPG--VLVDVSHQCRLOQDAYSAFCED--MDNVCHTLWC---SVGTTCHSKLDA- 510
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 496 WADGTFCGPHLCSGSCLEPEEVERPKFVVVDGMAPMGWGCSTCGGVQFSRECK 555
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 511 -VDGTRCGENKCKLSGECVPYG--FRPS-AVDGSGMSAMSI CSRS CGMGVQSAERQCT 566
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 556 DEEPONGERYCLGRPAKQSCHTBECPPDGKSFREQCEKKNAYNTDMDNLLQWPKY 615
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 567 QPTPKYKGRYCVGERKFRNLCLQCAPGRPSFRHVQCSHEDMLY---KGQLHTWVPV 623
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 616 AGVSPDRCKLFCRARGRSEFPVFEAKYIDGLCGPELTA--ICVRQCVKACDHYVDS 673
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 624 NDVNP---CELHCRANEYPAKLRDAVDDTFCYVARSRLCLINGICNVGCDPEIDS 680
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 674 PKLDKCYGCGKNSCKRVSSSLPT-NYGINDIVTIPAGATNIDVKQSHQPYQNDN 732
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 681 GAMEDRCGVCHNGSTCHTIVSGTEBAAGLYVDVGLIPAGAREIRIQE-----VAAAN 735
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 733 YLALXTADGO-YLALGNLAISAIEDILVKGTILKYSIATLERLOSFRPLPEPLTVOL 791
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 736 FLALNSEDPEKFLNGWTLIQ-WNGDVQVAGTFTYAR--GNWENLTSPPGTKEPWIQY 793
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 792 -----LTVPG-EVFPKXVYTFEVPN----- 811
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 794 PASRGPGGSGRGVPRFSTLHGSRPGSVSPGVTBPSBGPAAASTSVSPSLKPMNL 853
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 812 -----DYDFSMQSSKERATNTIIQ 830
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 854 VAAVRGQWGQAPLGLGWRRLVLMGRLLPTQLLFQESNPGVAYEYTIHREAGGHDEVP 913
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 831 PLHNAQVLTGDMSESSCCGAG--WQRETVACRD-PSGQ-----ASATONKALKPEDA 880
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 914 PVPFS-WAYGPWTKCLVTCGRGEEKGRHSPTCRGLVNSQGHNLQPLPAHCMAITGLEVCS 972
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 881 KP-----CESQ-----LCP 889
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Db 973 EPOFSICEMRLAIALCP 989

Search completed: May 7, 2004, 11:49:59  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:52:21 ; Search time 18 seconds  
(without alignments)  
2574.580 Million cell updates/sec

Title: US-09-989-687-4

Perfect score: 890  
Sequence: 1 MFPAAPAPRWLPFLILLILL.....CNKALKPDAPKCESQCLPL 890

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size: 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt 42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	890	100.0	890	1	Q9u799 homo sapien
2	35	3.9	905	1	AT58_HUMAN
3	12	1.3	205	1	AT58_MOUSE
4	12	1.3	630	1	AT58_YEAST
5	12	1.3	837	1	AT58_RAT
6	12	1.3	860	1	AT58_HUMAN
7	12	1.3	967	1	AT58_MOUSE
8	12	1.3	968	1	AT58_RAT
9	11	1.2	232	1	AT58_MOUSE
10	11	1.2	655	1	HGFA_MOUSE
11	10	1.1	152	1	NPW_PIG
12	10	1.1	165	1	NPW_HUMAN
13	10	1.1	185	1	NPW_MOUSE
14	10	1.1	245	1	AT58_HUMAN
15	10	1.1	245	1	AT58_MOUSE
16	10	1.1	245	1	AT58_RAT
17	10	1.1	480	1	SHBG_RABIT
18	10	1.1	492	1	PRTP_HUMAN
19	10	1.1	562	1	AT58_MOUSE
20	10	1.1	597	1	STIL_PANTR
21	10	1.1	741	1	PRO3_MOUSE
22	10	1.1	950	1	AT58_HUMAN
23	10	1.1	967	1	AT58_MOUSE
24	10	1.1	1007	1	CHC2_HUMAN
25	10	1.1	1013	1	PRPX_MACNE
26	10	1.1	1015	1	AT58_HUMAN
27	10	1.1	1081	1	CAML_FUGRU
28	10	1.1	1277	1	AT58_HUMAN
29	10	1.1	1935	1	AT58_HUMAN
30	10	1.1	2923	1	CIR2_HUMAN
31	9	1.0	107	1	PRK2_RAT
32	9	1.0	114	1	NPW_MOUSE
33	9	1.0	128	1	PRK2_MOUSE

34	9	1.0	131	1	OREX_HUMAN
35	9	1.0	131	1	OREX_PIG
36	9	1.0	133	1	YJ23_YEAST
37	9	1.0	134	1	SECR_RAT
38	9	1.0	151	1	PAX2_MOUSE
39	9	1.0	151	1	PAX2_RAT
40	9	1.0	157	1	RNS2_PANTR
41	9	1.0	161	1	UCN3_HUMAN
42	9	1.0	171	1	YCX7_YEAST
43	9	1.0	189	1	RMP2_MOUSE
44	9	1.0	197	1	PTCA_MOUSE
45	9	1.0	198	1	TR22_MOUSE

## ALIGNMENTS

RESULT 1  
AT58\_HUMAN STANDARD; PRT; 890 AA.  
AC Q9u799: Q9NZS0:  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE ADAMTS-8 precursor (BC 3.4.24.-) (A disintegrin and metalloproteinase  
DE with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2)  
DE (METH-8).  
GN ADAMTS8 OR METH2.  
OS Homo sapiens (Human).  
OC Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RT TISSUE=Lung;  
RX MEDLINE=99367466; PubMed=10438512;  
RA Valquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,  
RA Lombardo M., Irueña-Arispe M.L., and METH-2 are members of a new  
RT 'METH-1', a human ortholog of ADAMTS-1, and METH-2 are members of a new  
RT family of proteins with angio-inhibitory activity.";  
RT J. Biol. Chem. 274:23349-23357(1999).  
RN [2]  
RP SEQUENCE OF 195-440 FROM N.A.  
RX MEDLINE=20079168; PubMed=10610729;  
RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;  
RT "ADAM-TS8, a novel metalloproteinase of the ADAM-TS family located on  
RT mouse chromosome 9 and human chromosome 11.";  
RL Genomics 62:312-315(1999).  
CC -1- FUNCTION: Has anti-angiogenic properties.  
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
CC matrix (By similarity).  
CC -1- TISSUE SPECIFICITY: Highly expressed in adult and fetal lung, lower  
CC expression in brain, placenta, heart, stomach and fetal brain and  
CC kidney.  
CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important  
CC for a tight interaction with the extracellular matrix.  
CC -1- PFM: The precursor is cleaved by a furin endopeptidase (By  
CC similarity).  
CC -1- SIMILARITY: Belongs to peptidase family M28.  
CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
CC -1- SIMILARITY: Contains 2 TSP type-1 domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC EMBL: AF060153; AF048081.1; -  
CC EMBL: AF175283; AF25806.1; -

DR HSPB\_P34179; 11AG.  
 DR MEROPS; M12.226; -.  
 DR Genew; HGNC:224; ADAMTS8.  
 DR MIM; 605175; -.  
 DR GO; GO:0005179; F.integrin binding; TAS.  
 DR GO; GO:0008237; F.metalloproteinase activity; TAS.  
 DR GO; GO:0007345; P.embryogenesis and morphogenesis; TAS.  
 DR GO; GO:0008285; P.negative regulation of cell proliferation; TAS.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Regio1ysin; 1.  
 DR Pfam; PF00090; tsp1; 2.  
 DR PRINTS; PR01705; TSP1\_REPEAT.  
 DR SMART; SM00209; TSP1\_2.  
 DR PROSITE; PS00215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS00214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PS00092; TSP1; 2.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR Hydrolase; Metalloproteinase; zinc; signal; Glycoprotein; Zymogen;  
 KW Repeat; Extracellular matrix; Heparin-binding.  
 FT SIGNAL 1 27  
 FT PROPEP 28 214  
 FT CHAIN 215 890  
 FT DOMAIN 215 430  
 FT DOMAIN 439 526  
 FT DOMAIN 527 582  
 FT DOMAIN 584 690  
 FT DOMAIN 691 832  
 FT DOMAIN 834 889  
 FT DOMAIN 202 205  
 FT METAL 364 364  
 FT ACT\_SITE 365 365  
 FT METAL 368 368  
 FT METAL 374 374  
 FT CARBOHYD 345 345  
 FT CARBOHYD 401 401  
 FT CARBOHYD 466 466  
 FT CARBOHYD 491 491  
 FT CARBOHYD 600 600  
 FT CONFLICT 195 195  
 FT CONFLICT 413 440  
 SQ SEQUENCE 890 AA; 96671 MM; 57D70E03D5739D3 CRC64;  
 Query Match 100.0%; Score 890; DB 1; Length 890;  
 Best Local Similarity 100.0%; Pred No. 0;  
 Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 FNQPSRHPHYDTAILTRONFCQGEGLCDTLGVADIGTICDPNKSQSVTEDEGLAAH 360  
 DB 301 FNQPSRHPHYDTAILTRONFCQGEGLCDTLGVADIGTICDPNKSQSVTEDEGLAAH 360  
 QY 361 TLAEHLGVTHMPHDSKPCRTLFGPMGKHMAFLFVHLNQTLPWSPSCAMYTELLDG 420  
 DB 361 TLAEHLGVTHMPHDSKPCRTLFGPMGKHMAFLFVHLNQTLPWSPSCAMYTELLDG 420  
 QY 421 GHGDCILDPAGALPLPTGLFGMALYOLDQCRCQIFPDPFHGPNVSAQVCAQLMCHT 480  
 DB 421 GHGDCILDPAGALPLPTGLFGMALYOLDQCRCQIFPDPFHGPNVSAQVCAQLMCHT 480  
 QY 481 DGAEPICHTNGSLPWADGTPCGPHLCSGSCLEEEVERPKVVDGMAPMGMGEC 540  
 DB 481 DGAEPICHTNGSLPWADGTPCGPHLCSGSCLEEEVERPKVVDGMAPMGMGEC 540  
 QY 541 RTGGGVQSHRECPDREPONGRYCLARRAYQSCHEECPPOKSPREGQCKNAYN 600  
 DB 541 RTGGGVQSHRECPDREPONGRYCLARRAYQSCHEECPPOKSPREGQCKNAYN 600  
 QY 601 YTDMDGNLLQWPKYAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLGPEETLACV 660  
 DB 601 YTDMDGNLLQWPKYAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLGPEETLACV 660  
 QY 661 QCVKAGCDHVDSPRLDKGCVGCKGNSCRVSGSLPTNYGVNDIYTPAGATNIDYK 720  
 DB 661 QCVKAGCDHVDSPRLDKGCVGCKGNSCRVSGSLPTNYGVNDIYTPAGATNIDYK 720  
 QY 721 QSHPGVQNDGNVLAALKADGQYLLNGMLAISAEQDILVKTGSLATLERLOSF 780  
 DB 721 QSHPGVQNDGNVLAALKADGQYLLNGMLAISAEQDILVKTGSLATLERLOSF 780  
 QY 781 RLPEPLTQVLTLPGEVPPKVTTFPVNDVDPSSMSKERATNTIQLLPAQVVLG 840  
 DB 781 RLPEPLTQVLTLPGEVPPKVTTFPVNDVDPSSMSKERATNTIQLLPAQVVLG 840  
 QY 841 DMSECSSTGAGMORRTVECRDPGSGASATCNKALKPEDAPCESQICPL 890  
 DB 841 DMSECSSTGAGMORRTVECRDPGSGASATCNKALKPEDAPCESQICPL 890

RESULT 2  
 ID AT88\_MOUSE STANDARD; PRT; 905 AA.  
 AC P57110;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ADAMTS-8 precursor (BC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2).  
 GN ADAMTS8.  
 OS Mus musculus (mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_TaxID=10090;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20079168; PubMed=10610729;  
 RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;  
 RT "ADAM-TS8, a novel metalloproteinase of the ADAM-TS family located on mouse chromosome 9 and human chromosome 11.";  
 RL Genomics 62:312-315(1999)  
 CC -1- FUNCTION: Has anti-angiogenic properties (By similarity).  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed specifically in adult lung and heart and low expression during mouse development.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.  
 CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).



```

CC -1- SIMILARITY: Belongs to peptidase family M12B.
CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
CC -1- SIMILARITY: Contains 2 TSP type-1 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF15282; AAF25805.1; -.
CC HSSP; P34179; 11AG.
CC MEROPS; M12.226; -.
CC MGD; MGI:1353468; Adamts8.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR006025; Pept_M2n_BS.
CC InterPro; IPR001590; Peptidase_M2B.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP1.
CC Pfam; PF01421; Reprolysin; 1.
CC Pfam; PF00090; TSP_1; 2.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00209; TSP1; 2.
CC PROSITE; PS50215; ADAM_MERO; 1.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS00492; TSP1_2.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
CC Repeat; Extracellular matrix; Heparin-binding.
CC SIGNAL 1 28
CC PROPEP 29 228
CC CHAIN 229 905
CC DOMAIN 229 452
CC DOMAIN 453 541
CC DOMAIN 542 597
CC DOMAIN 599 705
CC DOMAIN 706 847
CC DOMAIN 848 904
CC METAL 378 378
CC ACT SITE 379 379
CC METAL 382 382
CC METAL 388 388
CC CARBOHYD 415 415
CC CARBOHYD 480 480
CC CARBOHYD 506 506
CC CARBOHYD 615 615
CC SEQUENCE 905 AA; 98879 MW; 124D4132B33A0CAE CRC64;

Query Match 3.9%; Score 35; DB 1; Length 905;
Best Local Similarity 100.0%; Pred. NO. 2.3e-22;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288c;
RX MEDLINE=84378720; PubMed=8091859;
RA Vanderbol M., Bolle P.-A., Dion C., Portetelle D., Hilger F.;
RT "DNA sequencing of a 36.2 kb fragment located between the PstI and
RL LAP loci of chromosome XI of Saccharomyces cerevisiae.";
Yeast 10:535-540(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z26877; CA81507.1; -.
CC EMBL; Z28146; CA81987.1; -.
CC PIR; S37804; S37804.
CC GenBank; 139903; -.
CC SGD; S0001630; YKL147C.
CC Hypothetical protein.
CC SEQUENCE 205 AA; 22673 MW; 01F162B361D9DCE CRC64;

Query Match 1.3%; Score 12; DB 1; Length 205;
Best Local Similarity 100.0%; Pred. NO. 0.0078;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 13 FILLILLILLPL 24
73 FILLILLILLPL 84

RESULT 4
ID ATSA_RAT STANDARD; PRT; 630 AA.
AC Q9ESP7; Q9ESP6; Q9ESP8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-4 precursor (EC 3.4.24.82) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1)
DE (Fragment).
CN ADAMTS4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M2star; TISSUE=Brain;
RX MEDLINE=20415831; PubMed=10961658;
RA Satch K., Suzuki N., Yokota H.;
RT "ADAMTS-4 (a disintegrin and metalloproteinase with thrombospondin
RT motifs) is transcriptionally induced in beta-amyloid treated rat
RT astrocytes.";
RL Neurosci. Lett. 289:177-180(2000).
CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
CC involved in its turnover. May play an important role in the
CC destruction of aggrecan in arthritic diseases. Cleaves aggrecan at
CC the 392-glu-|-Ala-393 site.
CC -1- CATALYTIC ACTIVITY: Glutamy| endopeptidase; bonds cleaved include
CC 370-Thr-Glu-Gly-Glu-|-Ala-His-Gly-Ser-377 in the interglobular
CC domain of mammalian aggrecan.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -1- TISSUE SPECIFICITY: Only expressed in brain.
CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -1- PTM: The precursor is cleaved by a furin endopeptidase.
CC -1- SIMILARITY: Belongs to peptidase family M12B.

```

```

CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB042272; BAB16474.1; -
DR EMBL; AB042273; BAB16475.1; -
DR EMBL; AB042274; BAB16476.1; -
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006025; Pept M Zn BS.
DR InterPro; IPR000884; Peptidase_M1B.
DR InterPro; IPR008085; TSP_1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00050; TSP_1; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50215; ADAM NEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
DR PROSITE; PS50214; DISINTEGRIN_2; FALSE NEG.
DR PROSITE; PS50092; TSP1; 1.
DR PROSITE; PS0142; ZINC_PROTEASE; 1.
KW Hydrolyase; Metalloprotease; Zinc; Glycoprotein; Zymogen;
KW Extracellular matrix.
FT NON TER 1 1
FT PROPEP 1 5
FT CHAIN 1 630
FT DOMAIN 6 232
FT DOMAIN 233 303
FT DOMAIN 313 368
FT DOMAIN 369 478
FT DOMAIN 479 630
FT DOMAIN 40 45
FT METAL 154 154
FT ACT SITE 155 155
FT METAL 158 158
FT METAL 164 164
FT CARBOHYD 96 96
FT CARBOHYD 474 474
SQ SEQUENCE 630 AA; 68384 MW; 634428753167C7EF CRC64;

Query Match 1.3%; Score 12; DB 1; Length 630;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 CSRTCGGVQFS 550
DB 325 CSRTCGGVQFS 336

RESULT 5
AT54 HUMAN STANDARD; PRT; 837 AA.
AC 075173; O9JUN83; -
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-4 precursor (EC 3.4.24.82) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1)
DE (ADMP-1).
GN ADAMTS4 OR KIAA0688.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.

Tissue=Brain;
RX MEDLINE=98403860; PubMed=9734811;
RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
RA Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. X.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro."
RL DNA Res. 5:169-176(1998).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=99286303; PubMed=10356395;
RA Torroella M.D., Burn T.C., Pratta M.A., Abbezzade I., Hollis J.M.,
RA Liu R.-Q., Rosenthal S.A., Copeland R.A., Decicco C.F., Wyman R.,
RA Rockwell A., Yang F., Duke J.L., Solomon K., George H., Bruckner R.,
RA Nagase H., Itoh Y., Ellis D.M., Ross H., Wiswall B.H., Murphy K.,
RA Hillman M.C. Jr., Hollis G.F., Newton R.C., Magolda R.L.,
RA Trzaskos J.M., Arner E.C.;
RT "Purification and cloning of aggrecanase-1: a member of the ADAMTS
RT family of proteases."
RL Science 284:1664-1666(1999).
RN [3]
RN SEQUENCE FROM N.A.
RX "Sawaji Y., Nagase H., Sakatvala J., Clark A.R.;
RX "ADAMTS-4 genomic locus."
RX Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RN PARTIAL SEQUENCE AND CHARACTERIZATION.
RX MEDLINE=20400518; PubMed=10827174;
RA Torroella M.D., Pratta M.A., Liu R.-Q., Abbezzade I., Ross H.,
RA Burn T.C., Arner E.C.;
RT "The thrombospondin motif of aggrecanase-1 (ADAMTS-4) is critical for
RT aggrecan substrate recognition and cleavage."
RL J. Biol. Chem. 275:25791-25797(2000).
CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
CC involved in its turnover. May play an important role in the
CC destruction of aggrecan in arthritic diseases. Could also be a
CC critical factor in the exacerbation of neurodegeneration in 333
CC Alzheimer's disease. Cleaves aggrecan at the 392-Glu-Ala-393
CC site.
CC -1- CATALYTIC ACTIVITY: Glutanyl endopeptidase; bonds cleaved include
CC 370-Thr-Glu-Gly-Glu-Ala-Arg-Ser-377 in the interglobular
CC domain of mammalian aggrecan.
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed in brain, lung and heart. Expressed
CC at very low level in placenta and skeletal muscles.
CC -1- INDUCTION: By interleukin-1.
CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -1- PTM: The precursor is cleaved by a furin endopeptidase.
CC -1- SIMILARITY: Belongs to peptidase family M2B.
CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
CC -1- SIMILARITY: Contains 1 TSP type-1 domain.
CC -1- CAUTION: Has sometimes been referred to as ADAMTS2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB014588; BAA31663.1; -
DR EMBL; AF148213; AAD41494.1; -
DR EMBL; AY044847; AAL02262.1; -
DR PIR; T00355; T00355.
DR MEROPS; M2.221; -.
DR GeneW; HGNC:220; ADAMTS4.
DR MIM; 603876; -
DR GO; GO:0008237; F:metallopeptidase activity; TAS.
DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.

```

DR GO:0001501; P.skeletal development; TAS.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006025; Pept\_M2n\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01421; Repolysin; 1.  
 DR Pfam; PF00090; tsp.1; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM0209; TSP1; 1.  
 DR PROSITE; PS50215; ADAM\_MPRO; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE; PS50092; TSP1; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KW Extracellular matrix.  
 FT SIGNAL 1 51 POTENTIAL.  
 FT PROPEP 52 212  
 FT CHAIN 213 837 ADAMTS-4.  
 FT DOMAIN 213 436 METALLOPROTEASE.  
 FT DOMAIN 437 519 DISINTEGRIN-LIKE.  
 FT DOMAIN 520 575 TSP TYPE-1.  
 FT DOMAIN 577 685 CYS-RICH.  
 FT DOMAIN 686 837 SPACER.  
 FT DOMAIN 247 252 POLY-ALA.  
 FT SITE 194 194 CYSTEINE SWITCH (POTENTIAL).  
 FT METAL 361 361 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT SITE 362 362 BY SIMILARITY.  
 FT METAL 365 365 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 371 371 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT CARBOHD 68 68 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHD 77 77 A -> T (IN REF. 1).  
 FT CONFLICT 626 626 R -> Q (IN REF. 3).  
 FT CONFLICT 682 682 G -> R (IN REF. 3).  
 SQ SEQUENCE 837 AA; 90224 MW; 5DF9C9AC137DF41F CRC64;

Query Match 1.3%; Score 12; DB 1; Length 837;  
 Best Local Similarity 100.0%; Pred. No. 0.024;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 CSRTCGGAGVGS 550  
 DB 532 CSRTCGGAGVGS 543

RESULT 6  
 ATIS6\_HUMAN STANDARD; PRT; 860 AA.  
 ID ATIS6\_HUMAN  
 AC Q9URF5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ADAMTS-6 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 6) (ADAM-TS 6) (ADAM-TS6).  
 GN ADAMTS6.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9395124; PubMed=10464288;  
 RA Hukariainen T.U., Hirohata S., Seidlin M.F., Agre S.S.;  
 RT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of  
 RT zinc metalloproteases";  
 RL J. Biol. Chem. 274:25555-25563 (1999).  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (By similarity).  
 CC -1- TISSUE SPECIFICITY: Expressed at low levels in placenta and barely  
 CC detectable in a number of other tissues.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important

CC for a tight interaction with the extracellular matrix.  
 CC -1- PFM: The precursor is cleaved by a furin endopeptidase (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 2 TSP type-1 domains.  
 CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF140674; MAD56357.1; -.  
 CC HSSP; P34179; 11AG.  
 CC MEROPS; M12.230; -.  
 CC Genew; HGNC:222; ADAMTS6.  
 CC MIM; 605008; -.  
 DR GO:0008237; F:metalloproteinase activity; TAS.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR006025; Pept\_M2n\_BS.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR002870; Peptidase\_M12B\_N.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01562; Pep\_M12B\_Prophep; 1.  
 DR Pfam; PF00090; Repolysin; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM0209; TSP1; 1.  
 DR PROSITE; PS50215; ADAM\_MPRO; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE; PS50092; TSP1; 1.  
 DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE; PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
 KW Repeat; Extracellular matrix.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT PROPEP 22 244 BY SIMILARITY.  
 FT CHAIN 245 452 ADAMTS-6.  
 FT DOMAIN 245 509 METALLOPROTEASE.  
 FT DOMAIN 453 509 DISINTEGRIN-LIKE.  
 FT DOMAIN 510 565 TSP TYPE-1.1.  
 FT DOMAIN 567 668 CYS-RICH.  
 FT DOMAIN 669 795 SPACER.  
 FT DOMAIN 796 852 TSP TYPE-1.2.  
 FT DOMAIN 852 852 POLY-ARG.  
 FT DOMAIN 68 71 POLY-GLY.  
 FT METAL 403 403 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT SITE 404 404 BY SIMILARITY.  
 FT METAL 407 407 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 413 413 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT CARBOHD 99 99 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHD 172 172 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHD 222 222 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHD 234 234 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHD 234 234 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHD 676 676 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHD 843 843 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 860 AA; 97098 MW; E57213015DEC82C5 CRC64;

Query Match 1.3%; Score 12; DB 1; Length 860;  
 Best Local Similarity 100.0%; Pred. No. 0.025;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGAV 547  
 DB 519 WGECSRTCGGAV 530

RESULT 7  
 ATIS1\_RAT

ID ATSL PAT STANDARD; PRT; 967 AA.  
 AC 09M001: 09M001; (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).  
 GN ADAMTS1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NC NCB1\_taxonomy=10116;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;  
 RA Liu X., Tu Y., Yin T., Johnstone E.M., Stephenson D.T., Clemens J.A.,  
 RA Little S.P.;  
 RT "Induction of a disintegrin and metalloproteinase with the  
 RT thrombospondin type I motif (ADAMTS).";  
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE OF 18-967 FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;  
 RX MEDLINE=20304099; PubMed=10847486;  
 RA Diamantis I., Luechi M., Hoessli M., Reichen J.;  
 RT "Cloning of the rat ADAMTS-1 gene and its down regulation in  
 RT endothelial cells in cirrhotic rats.";  
 RL Liver 20:165-172(2000).  
 CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
 CC involved in its turnover. Has angiogenic inhibitor activity (By  
 CC similarity). Active metalloproteinase, which may be associated with  
 CC various inflammatory processes as well as development of cancer  
 CC cachexia. May play a critical role in follicular rupture (By  
 CC similarity).  
 CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1683-Glu-|-Leu-1684  
 CC site, within the chondroitin sulfate attachment domain.  
 CC -1- CORRECTOR: Binds 1 zinc ion per subunit (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
 CC matrix (By similarity).  
 CC -1- INDUCTION: Down-regulated in endothelial cells derived from  
 CC cirrhotic liver.  
 CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important  
 CC for a tight interaction with the extracellular matrix.  
 CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By  
 CC similarity).  
 CC -1- SIMILARITY: Belongs to peptidase family M12B.  
 CC -1- SIMILARITY: Contains 1 disintegrin-like domain.  
 CC -1- SIMILARITY: Contains 3 TSP type-1 domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF149118; AAD34012.1; -  
 CC EMBL: AF304446; AAG29823.1; -  
 DR MEROPS: M12.222; -  
 DR InterPro: IPR006586; ADAM\_cysteine.  
 DR InterPro: IPR001762; Disintegrin.  
 DR InterPro: IPR006025; Pept\_M2n\_BS.  
 DR InterPro: IPR001590; Peptidase\_M12B.  
 DR InterPro: IPR002870; Peptidase\_M12B.  
 DR InterPro: IPR008084; TSP1.  
 DR InterPro: IPR008085; TSP1.  
 DR Pfam: PF01562; Pep\_M12B\_Dipep; 1.  
 DR Pfam: PF01421; RepProlysin; 1.  
 DR Pfam: PF00090; TSP1; 3.  
 DR PRINTS: PR01705; TSP1REPEAT.  
 DR SMART: SM00608; ACR; 1.  
 DR SMART: SM00209; TSP1; 3.

DR PROSITE: PS50215; ADAM\_MEROPS; 1.  
 DR PROSITE: PS00427; DISINTEGRIN\_1; FALSE\_NEG.  
 DR PROSITE: PS50214; DISINTEGRIN\_2; FALSE\_NEG.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 DR PROSITE: PS50092; TSP1; 3.  
 KW Hydrolase, Metalloproteinase; Zinc; Signal; Glycoprotein; Zymogen;  
 KW Repeat; Extracellular matrix; Heparin-binding.  
 FT SIGNAL 1 54  
 FT PROPEP 55 252  
 FT CHAIN 253 967  
 FT DOMAIN 253 475  
 FT DOMAIN 476 558  
 FT DOMAIN 559 614  
 FT DOMAIN 616 724  
 FT DOMAIN 725 857  
 FT DOMAIN 854 910  
 FT DOMAIN 911 967  
 FT DOMAIN 194 198  
 FT SITE 205 205  
 FT METAL 401 401  
 FT ACT\_SITE 402 402  
 FT METAL 405 405  
 FT METAL 411 411  
 FT CARBOHYD 547 547  
 FT CARBOHYD 720 720  
 FT CARBOHYD 764 764  
 FT CARBOHYD 782 782  
 FT CARBOHYD 945 945  
 FT CONFLICT 21 21  
 FT CONFLICT 26 31  
 FT CONFLICT 46 49  
 FT CONFLICT 72 72  
 FT CONFLICT 79 79  
 FT CONFLICT 249 249  
 FT CONFLICT 262 265  
 FT CONFLICT 607 607  
 FT CONFLICT 936 936  
 FT CONFLICT 962 962  
 SQ SEQUENCE 967 AA; 105705 MW; P93C86F6CDB4CF CRC64;  
 Query Match 1.34; Score 12; DB 1; Length 967;  
 Best Local Similarity 100.0%; Pred. No. 0.027;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 706 DIVITPAGATNI 717  
 DB 740 DIVITPAGATNI 751  
 RESULT 8  
 ID ATSL MOUSE STANDARD; PRT; 968 AA.  
 AC P97857: 054768;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase  
 with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).  
 GN ADAMTS1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCB1\_taxonomy=10090;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/SVJ;  
 RX MEDLINE=98110583; PubMed=9441751;  
 RA Kuno K., Iizasa H., Ohno S., Matsushima K.;  
 RT "The exon/intron organization and chromosome mapping of the mouse  
 RT ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";  
 RL Genomics 46:466-471(1997).  
 RN (2)  
 RP SEQUENCE FROM N.A.

RX MEDLINE=97150761; PubMed=8995297;  
RA Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,  
RA Matsushima K.;  
RT "Molecular cloning of a gene encoding a new type of metalloproteinase-  
RT disintegrin family protein with thrombospondin motifs as an  
RT inflammation associated gene.";  
RL J. Biol. Chem. 272:556-562(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Limb, and Mammary gland;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinci P., Prange C.,  
RA Rana S.S., Loguella N.A., Peters G.J., Adamson R.D., Mullany S.J.,  
RA Bosak S.A., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,  
RA Richards S., Wotley K.C., Walker K.J., Malek J.A., Gunaratne P.H.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [4]  
RP CHARACTERIZATION AND MUTAGENESIS OF GLU-403.  
RX MEDLINE=99303657; PubMed=10373500;  
RA Kuno K., Terahashima Y., Matsushima K.;  
RT "ADAMTS-1 is an active metalloproteinase associated with the  
RT extracellular matrix.";  
RL J. Biol. Chem. 274:18821-18826(1999).  
RN [5]  
RP FUNCTION.  
RX MEDLINE=20389568; PubMed=10930576;  
RA Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M.,  
RA Ohno H., Matsushima K.;  
RT "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.";  
RL FEBS Lett. 478:241-245(2000).  
RN [6]  
RP FUNCTION, AND INDUCTION.  
RX MEDLINE=20243757; PubMed=10781075;  
RA Robber R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,  
RA Richards U.S.;  
RT "Progestosterone-regulated genes in the ovulation process: ADAMTS-1 and  
RT cathepsin L proteases.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).  
CC -1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be  
CC involved in its turnover. Has antigenic inhibitor activity (By  
CC similarity). Active metalloproteinase, which may be associated with  
CC various inflammatory processes as well as development of cancer  
CC cachexia. May play a critical role in follicular rupture (By  
CC similarity).  
CC -1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1691-Glu-1692  
CC site, within the chondroitin sulfate attachment domain.  
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).  
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular  
CC matrix.  
CC -1- INDUCTION: Induced in vitro in colon adenocarcinoma cells by  
CC interleukin-1, or in vivo in kidney and heart by  
CC lipopolysaccharide. Also induced by LH stimulation in granulosa  
CC cells of preovulatory follicles.  
CC -1- DOMAIN: The spacer domain and the TSP type-1 domain are important  
CC for a tight interaction with the extracellular matrix.  
CC -1- PFM: The precursor is cleaved by a furin endopeptidase.  
CC -1- SIMILARITY: Belongs to peptidase family M12B.  
CC -1- SIMILARITY: Contains 1 disintegrin-like domain.

CC -1- SIMILARITY: Contains 3 TSP type-1 domain.  
CC -1- CAUTION: Ref.2 sequence differs from that shown due to a  
CC frameshift in position 7.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC DR EMBL; AB001735; BA24501.1; ALT INIT.  
CC DR EMBL; D67076; BA11088.1; ALT FRAME.  
CC DR EMBL; BC040382; AA40382.1; -  
CC DR EMBL; BC050834; AA50834.1; -  
CC DR MEROPS; M12.222; -  
CC DR MGD; MGI:109249; Adamts1.  
CC DR InterPro; IPR006586; ADAM\_cysteine.  
CC DR InterPro; IPR001762; Disintegrin.  
CC DR InterPro; IPR006025; Pept\_M\_zn\_BS.  
CC DR InterPro; IPR001590; Peptidase\_M12B.  
CC DR InterPro; IPR002870; Peptidase\_M12B\_N.  
CC DR InterPro; IPR000884; TSP1.  
CC DR InterPro; IPR008085; TSP1.  
CC DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
CC DR Pfam; PF01421; Reprolysin; 1.  
CC DR Pfam; PF00090; Tsp\_1; 3.  
CC DR PRINTS; PR01705; TSP1REPEAT.  
CC DR SMART; SM00608; ACR; 1.  
CC DR SMART; SM00209; TSP1; 3.  
CC DR PROSITE; PS0215; ADAM\_MEROP; 1.  
CC DR PROSITE; PS00427; DISINTEGRIN\_1; FALSE NEG.  
CC DR PROSITE; PS0214; DISINTEGRIN\_2; FALSE NEG.  
CC DR PROSITE; PS50092; TSP1; 3.  
CC DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
CC DR Hydrolase; Metalloproteinase; Zinc; Signal; Glycoprotein; Zymogen;  
CC KW Repeat; Extracellular matrix; Heparin-binding.  
CC FT SIGNAL; 1; 48  
CC FT PROPEP; 49; 253  
CC FT CHAIN; 254; 968  
CC FT DOMAIN; 254; 476  
CC FT DOMAIN; 477; 659  
CC FT DOMAIN; 560; 615  
CC FT DOMAIN; 618; 725  
CC FT DOMAIN; 726; 850  
CC FT DOMAIN; 855; 911  
CC FT DOMAIN; 912; 968  
CC FT DOMAIN; 195; 199  
CC FT SITE; 206; 206  
CC FT METIL; 402; 402  
CC FT ACT SITE; 403; 403  
CC FT METIL; 406; 406  
CC FT METIL; 412; 412  
CC FT CARBOHYD; 548; 548  
CC FT CARBOHYD; 721; 721  
CC FT CARBOHYD; 765; 765  
CC FT CARBOHYD; 783; 783  
CC FT CARBOHYD; 946; 946  
CC FT MUTAGEN; 403; 403  
CC FT CONFLICT; 335; 335  
CC FT CONFLICT; 425; 425  
CC SQ SEQUENCE; 968 AA; 105841 MW; 42EBDA55499F6C1 CRC64;  
CC  
CC Query Match 1.3%; Score 12; DB 1; Length 968;  
CC Best Local Similarity 100.0%; Pred. No. 0.027; 0; Indels 0; Gaps 0;  
CC Matches 12; Conservative 0; Mismatches 0;  
CC  
CC QY 706 DIVTIPAGATNI 717  
CC |||||  
CC DB 741 DIVTIPAGATNI 752

```

RESULT 9
FL3L MOUSE
ID FL3L MOUSE STANDARD; PRT; 232 AA.
AC P49772; Q64085;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE SL cytokine precursor (Fms-related tyrosine kinase 3 ligand) (Flt3
ligand) (Flt3L).
GN FLT3 OR FLT3L.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=94195428; PubMed=8145851;
RA Hannum C., Culpepper J., Campbell D., McClanahan T., Zurawski S.,
Bazan J.F., Kastelein R., Hudak S., Wagner J., Mattson J., Luh J.,
Duda G., Martini N., Peterson D., Menon S., Shanafelt A.,
Muench M., Keiser G., Nankawa R., Kennick D., Roncarolo M.G.,
Zlotnik A., Rosner O., Dibreuil P., Birnbaum D., Lee P.;
"Ligand for FLT3/FLK2 receptor tyrosine kinase regulates growth of
hematopoietic stem cells and is encoded by variant RNAs."
RL Nature 368:643-648(1994).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=SL/J;
RX MEDLINE=94084791; PubMed=7505204;
RA Lyman S.D., James L., Vanden Bos T., Devries P., Brasel K.,
Gliniak B., Hollingsworth L.T., Picha K.S., McKenna H.J.,
Splet R.R., Fletcher P.A., Maraskovsky E., Farrar T.,
Foxworth D., Williams D.E., Beckmann M.P.;
"Cloning of a ligand for the flt3/flk-2 tyrosine kinase
receptor: a proliferative factor for primitive hematopoietic cells."
RL Cell 75:1157-1167(1993).
RN (3)
RP SEQUENCE FROM N.A.; AND ALTERNATIVE SPLICING.
RX MEDLINE=96032581; PubMed=7566977;
RA Lyman S.D., Stocking K., Davison B., Fletcher F., Johnson L.,
Escobar S.;
"Structural analysis of human and murine flt3 ligand genomic loci."
RL Oncogene 11:1165-1172(1995).
RN (4)
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RX MEDLINE=95124710; PubMed=7824267;
RA Lyman S.D., James L., Escobar S., Downey H., de Vries P.,
Brasel K., Stocking K., Beckmann M.P., Copeland N.G.,
Cleveland L.S.;
"Identification of soluble and membrane-bound isoforms of the murine
flt3 ligand generated by alternative splicing of mRNAs."
RL Oncogene 10:149-157(1995).
RN (5)
RP SEQUENCE FROM N.A.
RA McClanahan T., Culpepper J., Campbell D., Wagner J.,
Franz-Bacon K., Mattson J., Tsai S., Luh J., Guimares M.J.,
Mattei M.-G., Rosnet O., Birnbaum D., Hannum C.;
Submitted (APR-1996) to the EMBL/Genbank/DBJ databases.
RL - FUNCTION: Stimulates the proliferation of early hematopoietic
cells. Synergizes well with a number of other colony stimulating
factors and interleukins.
CC - SUBUNIT: Homodimer (soluble isoform) (By similarity).
CC - SUBCELLULAR LOCATION: Type I membrane protein. Two soluble
isoforms are also produced by alternative splicing. One of which,
isoform 3/56, is biologically active, while the other, isoform
4/56beta1, is inactive.
CC - ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=4;
CC Name=1; Synonyms=6C;
CC IsoId=P49772-1; Sequence=Displayed;
CC Name=2; Synonyms=5H;
CC IsoId=P49772-2; Sequence=VSP_004253;
CC Name=3; Synonyms=B6;

```

```

CC CC IsoId=P49772-3; Sequence=VSP_004254, VSP_004255;
CC CC Name=4; Synonyms=B6Delta16;
CC CC IsoId=P49772-4; Sequence=VSP_004256;
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
-----
CC CC EMBL; U04807; AA18000.1; -
CC CC EMBL; L23636; AA39436.1; -
CC CC EMBL; U29875; AA90951.1; -
CC CC EMBL; U29875; AA90952.1; -
CC CC EMBL; S76459; AA33069.1; -
CC CC EMBL; S76461; AA33070.1; -
CC CC EMBL; S76464; AA33071.1; -
CC CC EMBL; U44024; AA33071.1; -
CC CC EMBL; U44024; AA3306.1; -
CC CC FIR; A49265; A49265.
CC CC MGD; MGI:95560; Flt3l.
CC CC InterPro; IPR004213; Flt3_lig.
CC CC Pfam; PF02947; flt3_lig; 1.
CC CC Cytokine; Glycoprotein; Transmembrane; Alternative splicing; Signal.
CC CC SIGNAL
CC CC CHAIN
CC CC DOMAIN 1 26
CC CC TRANSMEM 27 232
CC CC TRANSMEM 190 210
CC CC DOMAIN 211 232
CC CC DISULFID 31 112
CC CC DISULFID 71 156
CC CC DISULFID 120 161
CC CC CARBOHYD 127 127
CC CC CARBOHYD 152 152
CC CC VARSPLIC 164 232
CC CC FT FT 164 232
CC CC FT FT 169
CC CC FT FT 232
CC CC FT FT 163
CC CC FT FT 141
CC CC FT FT 198
CC CC FT FT 198
CC CC SQ SEQUENCE 232 AA; 26141 MW; 3A3680D3CB69FBA6 CRC64;
-----
Query Match 1.2%; Score 11; DB 1; Length 232;
Best local similarity 100.0%; Pared. No. 0.065;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
-----
Qy 14 LLLLLLLLLL 24
Db 190 LLLLLLLLLL 200
-----
RESULT 10
HGFA_HUMAN
ID HGFA_HUMAN STANDARD; PRT; 655 AA.
AC Q04756; Q14726;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hepatocyte growth factor activator precursor (EC 3.4.21.-) (HGF
activator) (HGFA).
GN HGFA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Liver, and Serum; PubMed=768365;  
 RX MEDLINE=93252878; PubMed=768365;  
 RA Miyazawa K., Shimomura T., Kitamura A., Kondo J., Morimoto Y.,  
 RA Kitamura N.;  
 RT "Molecular cloning and sequence analysis of the cDNA for a human  
 RT serine protease responsible for activation of hepatocyte growth  
 RT factor. Structural similarity of the protease precursor to blood  
 RT coagulation factor XII.";  
 RL J. Biol. Chem. 268:10024-10028 (1993).  
 RN [2]  
 RP SEQUENCE OF 40-655 FROM N.A.  
 RA Zhao S., Odell C.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1 FUNCTION: Activates hepatocyte growth factor (HGF) by  
 CC converting it from a single chain to a heterodimeric form.  
 CC -1 SUBUNIT: Dimer of a short chain and a long chain linked by a  
 CC disulfide bond.  
 CC -1 SUBCELLULAR LOCATION: Secreted as an inactive single-chain  
 CC precursor and is then activated to a heterodimeric form.  
 CC -1 TISSUE SPECIFICITY: Liver.  
 CC -1 SIMILARITY: Belongs to peptidase family S1.  
 CC -1 SIMILARITY: Contains 2 EGF-like domains.  
 CC -1 SIMILARITY: Contains 1 fibronectin type I domain.  
 CC -1 SIMILARITY: Contains 1 Kringle domain.  
 CC -1 CAUTION: It is uncertain whether Met-1 is the initiator.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC -----  
 DR EMBL: D14012; BAA03113.1;  
 DR EMBL: Z69923; -; NOT\_ANNOTATED\_CDS.  
 DR PIR: A46688; A46688.  
 DR HSRP: P00763; IDBO.  
 DR MEROPS: S01.228;  
 DR GeneW: HENC:4894; HGFPAC.  
 DR MIM: 604552;  
 DR GO: GO:0005576; C:extracellular; TAS.  
 DR GO: GO:0004252; F:serine-type endopeptidase activity; TAS.  
 DR GO: GO:0006508; P:proteolysis and peptidolysis; TAS.  
 DR InterPro: IPR000903; Cys\_Ser\_trypsin.  
 DR InterPro: IPR006209; EGF\_like.  
 DR InterPro: IPR000083; Fibrinctn.  
 DR InterPro: IPR000562; FN\_Type\_II.  
 DR InterPro: IPR006210; IEGF.  
 DR InterPro: IPR000001; Kringle.  
 DR InterPro: IPR001254; Peptidase\_S1.  
 DR InterPro: IPR001314; Peptidase\_S1A.  
 DR Pfam: PF00008; EGF; 2.  
 DR Pfam: PF00039; FN1; 1.  
 DR Pfam: PF00040; FN2; 1.  
 DR Pfam: PF00051; Kringle; 1.  
 DR Pfam: PF00089; trypsin; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PRINTS: PR00013; ENTPEPIL.  
 DR PRINTS: PR00018; KRINGLE.  
 DR ProDom: PD000995; FN\_Type\_II; 1.  
 DR ProDom: PD000395; Kringle; 1.  
 DR SMART: SM00181; EGF; 2.  
 DR SMART: SM00058; FN1; 1.  
 DR SMART: SM00059; FN2; 1.  
 DR SMART: SM00130; KR; 1.  
 DR SMART: SM00020; Tryp\_Spc; 1.

DR PROSITE: PS00022; EGF\_1; 2.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS00026; EGF\_3; 2.  
 DR PROSITE: PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE: PS00023; FIBRONECTIN\_2; 1.  
 DR PROSITE: PS00021; KRINGLE\_1; 1.  
 DR PROSITE: PS00070; KRINGLE\_2; 1.  
 DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 KW Hydroxylase; Glycoprotein; Plasma; Serine protease; Kringle; Signal;  
 KW EGF-like domain; Repeat; Zymogen.  
 FT SIGNAL 1 30  
 FT PROPEP 31 372  
 FT CHAIN 373 407  
 FT CHAIN 408 655  
 FT DOMAIN 108 148  
 FT DOMAIN 160 198  
 FT DOMAIN 200 240  
 FT DOMAIN 241 279  
 FT DOMAIN 286 367  
 FT DOMAIN 408 655  
 FT ACT\_SITE 447 447  
 FT ACT\_SITE 497 497  
 FT ACT\_SITE 598 598  
 FT DISULFID 108 133  
 FT DISULFID 122 148  
 FT DISULFID 164 175  
 FT DISULFID 169 186  
 FT DISULFID 188 197  
 FT DISULFID 202 230  
 FT DISULFID 228 237  
 FT DISULFID 245 256  
 FT DISULFID 250 267  
 FT DISULFID 278 278  
 FT DISULFID 286 367  
 FT DISULFID 307 349  
 FT DISULFID 338 362  
 FT DISULFID 394 521  
 FT DISULFID 432 448  
 FT DISULFID 440 510  
 FT DISULFID 535 604  
 FT DISULFID 567 583  
 FT DISULFID 594 622  
 FT CARBOHYD 48 48  
 FT CARBOHYD 290 290  
 FT CARBOHYD 468 468  
 FT CARBOHYD 492 492  
 FT CARBOHYD 546 546  
 FT CONFLICT 644 644  
 SQ SEQUENCE 655 AA; 70681 MW; 2CF72F1E1B86ED7 CR664;  
 Query Match 1.2%; Score 11; DB 1; Length 655;  
 Best Local Similarity 100.0%; Pred. No. 0.15;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 12 PFIILILILIL 22  
 18 PFIILILILIL 28  
 RESULT 11  
 NPW\_PIG STANDARD; PRT; 152 AA.  
 AC OSW135;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Neuropeptide W precursor (Preproprotein W8) (PPI8) [Contains:  
 DE Neuropeptide W-23 (NPW23) (L8); Neuropeptide W-30 (NPW30) (L8C)].  
 GN NPW OR PPI8.

OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 OC NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A., SEQUENCE OF 33-52, AND SYNTHESIS OF NPW23 AND  
 RP NPW30.  
 RX MEDLINE=22229377; PubMed=12130646;  
 RA Shimomura Y., Harada M., Goto M., Sugo T., Matsumoto Y., Abe M.,  
 RA Watanabe T., Asami T., Kitada C., Mori M., Onda H., Fujino M.;  
 RT "Identification of neuropeptide W as the endogenous ligand for orphan  
 RT G-protein-coupled receptor GPR7 and GPR8.";  
 RL J. Biol. Chem. 277:35826-35832(2002).  
 CC -1- FUNCTION: Plays a regulatory role in the organization of  
 CC neuroendocrine signals accessing the anterior pituitary gland.  
 CC stimulates water drinking and food intake. May play a role in the  
 CC hypothalamic response to stress (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the neuropeptide B/W family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL, AB084277; BAC07173.1; -  
 KW Neuropeptide; Cleavage on pair of basic residues; Signal.  
 FT SIGNAL 1 32  
 FT PEPTIDE 33 55 NEUROPEPTIDE W-23.  
 FT PROPEP 65 152 NEUROPEPTIDE W-30.  
 FT PROPEP 152 16347 MW; 454284DD448067E4 CRC64;  
 SQ SEQUENCE 152 AA; 16347 MW; 454284DD448067E4 CRC64;  
 Query Match 1.1%; Score 10; DB 1; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 0.34; Mismatches 0; Indels 0;  
 Matches 10; Conservative 0; Mismatches 0;  
 QY 15 LLLLLLLPL 24  
 DB 19 LLLLLLLPL 28  
 RESULT 12  
 NPW\_HUMAN STANDARD; PRT; 165 AA.  
 AC Q8N729;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Neuropeptide W precursor (Preproprotein 18) (pPL8) [Contains:  
 DE Neuropeptide W-23 (NPW23) (hL6); Neuropeptide W-30 (NPW30) (hL8C)].  
 GN NPW OR PPNW OR PPL8.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND SYNTHESIS OF NPW23 AND  
 RP NPW30.  
 RX MEDLINE=22229377; PubMed=12130646;  
 RA Shimomura Y., Harada M., Goto M., Sugo T., Matsumoto Y., Abe M.,  
 RA Watanabe T., Asami T., Kitada C., Mori M., Onda H., Fujino M.;  
 RT "Identification of neuropeptide W as the endogenous ligand for orphan  
 RT G-protein-coupled receptors GPR7 and GPR8.";  
 RL J. Biol. Chem. 277:35826-35832(2002).  
 CC -1- FUNCTION: Plays a regulatory role in the organization of  
 CC neuroendocrine signals accessing the anterior pituitary gland.  
 CC stimulates water drinking and food intake. May play a role in the  
 CC hypothalamic response to stress (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: Belongs to the neuropeptide B/W family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL, AB084276; BAC07172.1; ALT\_INIT.  
 KW Neuropeptide; Cleavage on pair of basic residues; Signal.  
 FT SIGNAL 1 32  
 FT PEPTIDE 33 55 NEUROPEPTIDE W-23.  
 FT PROPEP 65 165 NEUROPEPTIDE W-30.  
 FT PROPEP 165 18048 MW; F426AF4B7F6B604 CRC64;  
 SQ SEQUENCE 165 AA; 18048 MW; F426AF4B7F6B604 CRC64;  
 Query Match 1.1%; Score 10; DB 1; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 0.36; Mismatches 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RT "Identification of natural ligands for the orphan G protein-coupled  
 RT receptors GPR7 and GPR8.";  
 RL J. Biol. Chem. 278:176-783(2003).  
 CC -1- FUNCTION: Plays a regulatory role in the organization of  
 CC neuroendocrine signals accessing the anterior pituitary gland.  
 CC stimulates water drinking and food intake. May play a role in the  
 CC hypothalamic response to stress (By similarity). NPW23 activates  
 CC GPR7 and GPR8 more efficiently than NPW30.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Detected at high levels in the substantia  
 CC nigra, fetal kidney and testes; at lower levels in testis,  
 CC uterus, ovary and placenta. Not detectable in many regions of the  
 CC central nervous system. Also detected at high levels in  
 CC lymphoblastic leukemia and colorectal adenocarcinoma.  
 CC -1- SIMILARITY: Belongs to the neuropeptide B/W family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL, AB084276; BAC07172.1; ALT\_INIT.  
 KW Neuropeptide; Cleavage on pair of basic residues; Signal.  
 FT SIGNAL 1 32  
 FT PEPTIDE 33 55 NEUROPEPTIDE W-23.  
 FT PROPEP 65 165 NEUROPEPTIDE W-30.  
 FT PROPEP 165 18048 MW; F426AF4B7F6B604 CRC64;  
 SQ SEQUENCE 165 AA; 18048 MW; F426AF4B7F6B604 CRC64;  
 Query Match 1.1%; Score 10; DB 1; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 0.36; Mismatches 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 15 LLLLLLLPL 24  
 DB 19 LLLLLLLPL 28  
 RESULT 13  
 NPW\_RAT STANDARD; PRT; 185 AA.  
 AC O8KTM5;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Neuropeptide W precursor (Preproprotein 18) (pPL8) [Contains:  
 DE Neuropeptide W-23 (NPW23) (L6); Neuropeptide W-30 (NPW30) (L8C)].  
 GN NPW OR PPL8.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RX MEDLINE=22229377; PubMed=12130646;  
 RA Shimomura Y., Harada M., Goto M., Sugo T., Matsumoto Y., Abe M.,  
 RA Watanabe T., Asami T., Kitada C., Mori M., Onda H., Fujino M.;  
 RT "Identification of neuropeptide W as the endogenous ligand for orphan  
 RT G-protein-coupled receptors GPR7 and GPR8.";  
 RL J. Biol. Chem. 277:35826-35832(2002).  
 CC -1- FUNCTION: Plays a regulatory role in the organization of  
 CC neuroendocrine signals accessing the anterior pituitary gland.  
 CC stimulates water drinking and food intake. May play a role in the  
 CC hypothalamic response to stress (By similarity). NPW23 activates  
 CC GPR7 and GPR8 more efficiently than NPW30.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Detected at high levels in the substantia  
 CC nigra, fetal kidney and testes; at lower levels in testis,  
 CC uterus, ovary and placenta. Not detectable in many regions of the  
 CC central nervous system. Also detected at high levels in  
 CC lymphoblastic leukemia and colorectal adenocarcinoma.  
 CC -1- SIMILARITY: Belongs to the neuropeptide B/W family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL, AB084276; BAC07172.1; ALT\_INIT.  
 KW Neuropeptide; Cleavage on pair of basic residues; Signal.  
 FT SIGNAL 1 32  
 FT PEPTIDE 33 55 NEUROPEPTIDE W-23.  
 FT PROPEP 65 165 NEUROPEPTIDE W-30.  
 FT PROPEP 165 18048 MW; F426AF4B7F6B604 CRC64;  
 SQ SEQUENCE 165 AA; 18048 MW; F426AF4B7F6B604 CRC64;  
 Query Match 1.1%; Score 10; DB 1; Length 165;  
 Best Local Similarity 100.0%; Pred. No. 0.36; Mismatches 0;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Stimulates water drinking and food intake. May play a role in the hypothalamic response to stress. When injected into the lateral cerebromedial, it elevates prolactin (PR) and corticosterone and lowers growth hormone (GH) release.

-1- SUBCELLULAR LOCATION: Secreted.

-1- SIMILARITY: Belongs to the neuropeptide B/w family.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: AB084278; BAC07174.1; -

KW Neuropeptide; Cleavage on pair of basic residues; Signal.

FT SIGNAL 1 41

FT PEPTIDE 42 64 NEUROPEPTIDE W-23.

FT PPPTIDE 42 71 NEUROPEPTIDE W-30.

FT PROPER 74 185 POTENTIAL.

SO SEQUENCE 185 AA; 20231 MW; 2A4C82DFE0A8F6B9 CRC64;

Query Match 1.1%; Score 10; DB 1; Length 185;  
Best Local Similarity 100.0%; Pred.No. 0.4; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0;

QY 15 LLLLLLLLLL 24  
Db 28 LLLLLLLLLL 37

RESULT 14  
AT54 BOVIN STANDARD; PRT; 245 AA.

AC QPRT53; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE ADAMTS-4 (EC 3.4.24.82) (A disintegrin and metalloproteinase with chromospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1) (ADAMP-1) (Fragments).

GN ADAMTS4.

OS Bos taurus (Bovine); Chordata; Craniata; Vertebrata; Euteleostomi; OC Eukaryota; Metazoa; Chordata; Cephalochordata; Ruminantia; Pecora; Bovidae; OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE OF 18-227 FROM N.A.

RX MEDLINE=99333577; PubMed=10403768;

RA Flannery C.R., Little C.B., Hughes C.E., Cateson B.;

RT "Expression of ADAMTS homologues in articular cartilage.";

RL Biochem. Biophys. Res. Commun. 260:318-322(1999).

RN [2]

RP SEQUENCE OF 18-227 FROM N.A.

RX MEDLINE=20092827; PubMed=10625599;

RA Curtis C.L., Hughes C.E., Flannery C.R., Little C.B., Harwood J.L.,

RT "n-3 fatty acids specifically modulate catabolic factors involved in articular cartilage degradation.";

RL J. Biol. Chem. 275:721-724(2000).

RN [3]

RP SEQUENCE OF 1-27; 228-233 AND 234-245.

RX TISSUE=Cartilage; PubMed=10356395;

RX MEDLINE=99286303; PubMed=10356395;

RA Tortorella M.D., Burn T.C., Pratta M.A., Abasrade I., Hollis J.M.,

RA Liu R., Rosenfeld S.A., Copeland R.A., Decicco G.P., Wynn R.,

RA Rockwell A., Yang F., Duke J.L., Solomon K., George H., Bruckner R.,

RA Nagase H., Itoh Y., Ellis D.M., Ross H., Miewall B.H., Murphy K.,

RA Hillman M.C. Jr., Hollis G.F., Newton R.C., Magolda R.L.,

RA Tzavakos J.M., Arner E.C.;

RT "Purification and cloning of aggrecanase-1: a member of the ADAMTS

family of proteins.";

RL Science 284:1664-1666(1999).

-1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be involved in its turnover. May play an important role in the destruction of aggrecan in arthritic diseases. Cleaves aggrecan at the 392-Glu-Ala-393 site.

-1- CATALYTIC ACTIVITY: Glutamate endopeptidase; bonds cleaved include 370-Thr-Glu-Gly-Glu-Ala-Arg-Gly-Ser-377 in the interglobular domain of mammalian aggrecan.

-1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

-1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).

-1- INDUCTION: By interleukin-1.

-1- PM: The precursor is cleaved by a furin endopeptidase (By similarity).

-1- SIMILARITY: Belongs to peptidase family M12B.

-1- CAUTION: Has sometimes been referred to as ADAMTS2.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: AF192770; AA07176.1; -

DR HSR: P34179; AATG.

DR InterPro: IPR006025; Pept\_M\_Zn\_BS.

DR InterPro: IPR001590; Peptidase\_M12B.

DR PROSITE: PS00142; ZINC\_PROTEASE; 1.

DR PROSITE: PS50215; ADAM\_MPRO; 1.

KW Hydrolyase; Metalloproteinase; Zinc; Glycoprotein; Extracellular matrix.

FT NON\_TER 1 1

FT NON\_TER 27 28

FT NON\_TER 27 28

FT DOMAIN 27 28

FT METAL 57 57

FT METAL 57 57

FT ACT SITE 58 58

FT METAL 61 61

FT METAL 67 67

FT METAL 227 228

FT NON\_TER 233 234

FT NON\_TER 245 245

FT NON\_TER 245 245

SO SEQUENCE 245 AA; 26391 MW; 7C687968CAB431D CRC64;

Query Match 1.1%; Score 10; DB 1; Length 245;  
Best Local Similarity 100.0%; Pred.No. 0.5;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 PWAADGTPCGP 504  
Db 184 PWAADGTPCGP 193

RESULT 15  
C10C HUMAN STANDARD; PRT; 245 AA.

ID C10C\_HUMAN

AC P02747; Q96DL2; Q96H05;

DT 21-JUL-1986 (Rel. 01, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Complement C1q subcomponent, C chain precursor.

GN C1QG OR C1QC (Human).

OS Homo sapiens (Human); Chordata; Craniata; Vertebrata; Euteleostomi; OC Eukaryota; Metazoa; Chordata; Cephalochordata; Primates; Hominiidae; Homo.

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX TISSUE=Monocytes;

RX MEDLINE=91174759; PubMed=1706597;

RA Sellar G.C., Blake D.J., Reid K.B.M.;

RT "Characterization and organization of the genes encoding the A-, B-

RT and C-chains of human complement subcomponent C1q. The complete  
 RT derived amino acid sequence of human C1q.";  
 RL Biochem. J. 274:481-490(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Cerebellum;  
 RA Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y.,  
 RA Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,  
 RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,  
 RA Kawakami B., Nagai K., Isegai T., Sugano S.;  
 RA "NEO human CDNA sequencing project";  
 RT Submitted (Oct-2001) to the EMBL/Genbank/DBJ databases.  
 RL [3]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Brain;  
 RC MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters K.J., Abramson R.D., Muljany S.J.,  
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Pahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Nadeau A., Young A.C., Shevchenko I., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez Y.C., Grimwood J., Schmutz J., Myers R.W.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RA "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 29-122.  
 RC MEDLINE=80020137; PubMed=468087;  
 RA Reid K.B.M.;  
 RA "Complete amino acid sequences of the three collagen-like regions  
 RT present in subcomponent C1q of the first component of human  
 RL complement";  
 RL Biochem. J. 179:367-371(1979).  
 RN [5]  
 RP REVIEW OF C1Q DEFICIENCY.  
 RC MEDLINE=98450587; PubMed=9777412;  
 RA Petry F.;  
 RA "Molecular basis of hereditary C1q deficiency.";  
 RT Immunobiology 199:286-294(1998).  
 RL [6]  
 RP FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD  
 CC C1. THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE  
 CC COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT  
 CC C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1  
 CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE  
 CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.  
 CC -1- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R  
 CC AND S IN THE MOLAR RATION OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED  
 CC OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE  
 CC A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE  
 CC C CHAIN.  
 CC -1- PTM: C-linked glycans consist of Glc-Gal disaccharides bound to  
 CC the oxygen atom of post-translationally added hydroxyl groups.  
 CC -1- DISEASE: Defects in C1q are a cause of C1q deficiency  
 CC [MIM:120575]. It is a rare genetic disorder which is associated  
 CC with recurrent infections and a high prevalence of lupus  
 CC erythematosus-like symptoms. It is characterized by a loss of  
 CC activation of the complement classical pathway.  
 CC -1- SIMILARITY: Contains 1 C1q domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AK057792; BAB71575.1; -;  
 CC EMBL; BC009016; AA09016.1; -;  
 CC PIR; S14351; C1HUOC.  
 CC Gene; HGNC:1245; C1Q.  
 CC MIM; 120575; -;  
 CC GO; GO:0005576; C:extracellular; NAS.  
 CC GO; GO:0003793; F:defense/immunity protein activity; NAS.  
 CC GO; GO:0006955; P:immune response; NAS.  
 CC InterPro; IPR001073; C1q.  
 CC InterPro; IPR008160; Collagen.  
 CC InterPro; IPR008983; TNF\_like.  
 CC Pfam; PF00386; C1q; 1.  
 CC Pfam; PF03391; Collagen; 1.  
 CC PRINTS; PR00007; COMPLEMENTC1Q.  
 CC SMART; SM0110; C1Q; 1.  
 CC PROSITE; PS0113; C1q; 1.  
 CC Complement pathway; Plasma; Hydroxylation; glycoprotein; Collagen;  
 CC Repeat; Signal; Disease mutation.  
 CC SIGNAL 1 28  
 CC CHAIN 29 245  
 CC DOMAIN 31 112  
 CC DISULFID 32 32  
 CC MOD\_RES 36 36  
 CC MOD\_RES 39 39  
 CC MOD\_RES 42 42  
 CC MOD\_RES 45 45  
 CC MOD\_RES 54 54  
 CC MOD\_RES 57 57  
 CC MOD\_RES 63 63  
 CC MOD\_RES 66 66  
 CC MOD\_RES 71 71  
 CC MOD\_RES 75 75  
 CC CARBOHYD 81 81  
 CC MOD\_RES 84 84  
 CC CARBOHYD 84 84  
 CC MOD\_RES 93 93  
 CC MOD\_RES 96 96  
 CC MOD\_RES 99 99  
 CC MOD\_RES 105 105  
 CC VARIANT 43 43  
 CC CONFLICT 14 14  
 CC CONFLICT 23 23  
 CC CONFLICT 57 57  
 CC CONFLICT 66 66  
 CC CONFLICT 72 72  
 CC CONFLICT 84 84  
 CC CONFLICT 87 87  
 CC CONFLICT 90 90  
 CC CONFLICT 215 215  
 CC SEQUENCE 245 AA; 25774 MW; FAL7117EB7ABFC12 CRC64;  
 CC  
 CC Query Match 1.1%; Score 10; DB 1; Length 245;  
 CC Best Local Similarity 100.0%; Pred No. 0.5; Indels 0; Gaps 0;  
 CC Matches 10; Conservative 0; Mismatches 0;  
 CC  
 CC QY 15 LLLLLLLPL 24  
 CC DB 15 LLLLLLLPL 24

Search completed: May 7, 2004, 11:59:06  
 Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:42:50 ; Search time 65 Seconds  
(without alignments)  
3668.727 Million cell updates/sec

Title: US-09-989-687-4  
Perfect score: 4853  
Sequence: 1 MFPAAPAPRWLPFLLLLL.....CNKAKPEDAKPCESQCTPL 890

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp19808:\*  
2: Geneseqp19908:\*  
3: Geneseqp20008:\*  
4: Geneseqp20018:\*  
5: Geneseqp20028:\*  
6: Geneseqp20038:\*  
7: Geneseqp20038:\*  
8: Geneseqp20048:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4853	100.0	890	2	AAV49502 Human MET
2	4853	100.0	890	2	AAV49502 Human MET
3	4822	99.4	890	6	ABP96306 Human ADA
4	4807.5	99.1	889	4	AAV74946 Human ADA
5	3919.5	60.8	905	4	AAV73284 Human ADA
6	3713	76.5	680	3	AAV21251 Human met
7	3244	48.3	896	3	AAV21265 Human met
8	2344	48.3	896	6	ABU08387 Murine ma
9	2328.5	48.0	950	4	AAV49501 Human ADA
10	2328.5	48.0	950	4	AAV73549 Human ADA
11	2328.5	48.0	950	4	AAV50002 Human MET
12	2328.5	48.0	950	2	AAV80285 Human int
13	2328.5	48.0	968	4	AAV50011 Protein;
14	2327.5	48.0	967	2	AAV04142 Human Tan
15	2317.5	47.8	950	3	AAV53899 Human act
16	2315.5	47.7	967	2	AAV78189 Human sec
17	2315.5	47.7	967	6	AAV57139 Human sec
18	2315.5	47.7	967	6	AAV41003 Human sec
19	2315.5	47.7	967	6	AAV91631 Human sec
20	2315.5	47.7	967	7	AAV74267 Human sec
21	2315.5	47.7	967	7	AAV73948 Human sec
22	2280.5	47.0	481	2	AAV04145 Rat Tanco
23	2146.5	44.2	924	5	AAV70062 Human NOV
24	2138.5	44.1	950	5	AAV22541 Human pro
25	2138.5	44.1	950	6	AAV09520 Human pro

26	2135.5	44.0	950	4	AAV62299 Human met
27	2131.5	43.9	950	6	AAV40092 Human ADA
28	2115.5	43.6	952	5	AAV74751 Human pro
29	2103.5	43.3	727	2	AAV78435 Human ADA
30	2047.5	42.1	928	5	AAV72899 Human met
31	1985.5	41.0	823	6	AAV08383 Human met
32	1989	41.0	364	4	AAV09710 Human gen
33	1989	41.0	364	7	AAV22088 Human pro
34	1842.5	38.0	837	2	AAV75425 Human agv
35	1842.5	38.0	837	4	AAV78228 Human agv
36	1842.5	38.0	837	7	AAV85488 Human agv
37	1842.5	38.0	840	3	AAV21256 Human met
38	1841.5	37.9	837	3	AAV94429 Human pro
39	1841.5	37.9	837	4	AAV6178 Human pro
40	1841.5	37.9	837	4	AAV29199 Human pro
41	1841.5	37.9	837	6	AAV58575 Human pro
42	1841.5	37.9	837	6	AAV08123 Human pro
43	1841.5	37.9	837	6	AAV84438 Human sec
44	1841.5	37.9	837	6	AAV66312 Human sec
45	1841.5	37.9	837	6	AAV65702 Human sec

## ALIGNMENTS

RESULT 1  
AAV49502  
ID AAV49502 standard; protein, 890 AA.  
XX  
AC AAV49502;  
XX  
DT 10-JAN-2000 (first entry)  
XX  
DE Human MET/2 protein.  
XX  
KW Human; MET/1; MET/2; anti-angiogenic; metalloprotease thrombospondin;  
KW cancer; diagnosis; hyperproliferative disorder; autoimmune disease;  
KW angiogenesis inhibitor; abnormal wound healing; inflammation;  
KW rheumatoid arthritis; psoriasis; endometrial bleeding disorder;  
KW diabetic retinopathy; macula degeneration; haemangioma; detection;  
KW arterial-venous malformation; immune deficiency.  
XX  
OS Homo sapiens.  
XX  
PN MO9937660-A1.  
XX  
PD 29-JUL-1999.  
XX  
PF 22-JAN-1999; 99WO-US001313.  
XX  
PR 23-JAN-1998; 98US-0072288P.  
XX  
PR 28-AUG-1998; 98US-0098539P.  
XX  
PA (IRU//) IRU/AA-Arispe L.  
PA (HAST//) HASTINGS G A.  
PA (ROBE//) ROSEN S M.  
XX  
PI Irrela-Arispe L, Hastings GA, Ruben SM;  
XX  
DR WPI, 1999-590684/50.  
XX  
DR N-PSDB; AA232001.  
XX  
PT New isolated metalloprotease thrombospondin polypeptides, useful for  
PT treating hyperproliferative disorders, cancers or autoimmune disorders.  
XX  
PS Claim 10; Fig 2; 457pp; English.  
XX  
CC AA232000 and AA232001 encode, and AAV49501 and AAV49502 represent, human  
CC metalloprotease thrombospondin (MET/1) proteins MET/1 and MET/2  
CC respectively. MET/1 and MET/2 have been found to be potent inhibitors of  
CC angiogenesis both in vitro and in vivo. They can be used for treating  
CC cancer and other disorders related to angiogenesis including abnormal  
CC wound healing, inflammation, rheumatoid arthritis, psoriasis, endometrial

bleeding disorders, diabetic retinopathy, some forms of macula degeneration, haemangiomas, and arterial-venous malformations. They may be useful in treating deficiencies or disorders of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilisation (chemotaxis) of immune cells. The etiology of these immune deficiencies or disorders may be genetic, somatic, such as cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or toxins), or infectious. They can also be used to treat inflammatory conditions, both chronic and acute conditions. The products can also be used for detection and diagnosis. AA32002 to AA32080, and AA49503 to AA49511 represent sequences given in the exemplification of the present invention

Sequence 890 AA;

Query Match 100.0%; Score 4853; DB 2; Length 890;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPAAPAPRMPLPPLLLLLPLARAPAPAPAGASATVPTLPAGAGLALHLSA 60  
DB 1 MFPAAPAPRMPLPPLLLLLPLARAPAPAPAGASATVPTLPAGAGLALHLSA 60  
QY 61 FGKGFVRLAPDDSFLLAPEFKIERLGGSGRATGGERGLRGCFSGTVNGEPESLAAVSLC 120  
DB 61 FGKGFVRLAPDDSFLLAPEFKIERLGGSGRATGGERGLRGCFSGTVNGEPESLAAVSLC 120  
QY 121 RGLSGSFLLDGEFTTIOQAGSGLAOPHRLQWGPAPAPLRGPWEVETEGGQORER 180  
DB 121 RGLSGSFLLDGEFTTIOQAGSGLAOPHRLQWGPAPAPLRGPWEVETEGGQORER 180  
QY 181 GDHGEDSEESQEEAEAGASPPPLGTSRTKAFSEAFVETLVAADASMAAFYADL 240  
DB 181 GDHGEDSEESQEEAEAGASPPPLGTSRTKAFSEAFVETLVAADASMAAFYADL 240  
QY 241 QNHILITMSVAAIKYKPSINSINLMVYKYLVEDEKQBEVSDNGLLTNFCWQOR 300  
DB 241 QNHILITMSVAAIKYKPSINSINLMVYKYLVEDEKQBEVSDNGLLTNFCWQOR 300  
QY 301 FNPQSDRPHRYCTRAILLTRONFGQEGSLCPTLGVADIGTCDNKSCSYIEBEGLOAH 360  
DB 301 FNPQSDRPHRYCTRAILLTRONFGQEGSLCPTLGVADIGTCDNKSCSYIEBEGLOAH 360  
QY 361 TLAEHLGAVLSMHDSPKCTRLFGPMKHVMAPLFVHNLQTLPMSPCSAMYTELLDG 420  
DB 361 TLAEHLGAVLSMHDSPKCTRLFGPMKHVMAPLFVHNLQTLPMSPCSAMYTELLDG 420  
QY 421 GHGDCILAPGAALPLPTGLGPMALYQLDOQCQIIEBDRPHCPNTSAOVCAQLMCHT 480  
DB 421 GHGDCILAPGAALPLPTGLGPMALYQLDOQCQIIEBDRPHCPNTSAOVCAQLMCHT 480  
QY 481 DGAEPILCTKNGSLPMADGTPCGPGLCSGSCULPEEVEVERPKYVDGMAPMPWMECS 540  
DB 481 DGAEPILCTKNGSLPMADGTPCGPGLCSGSCULPEEVEVERPKYVDGMAPMPWMECS 540  
QY 541 RTGGGVOFSHRECDPEPONGRCYCLGRAPKQSCHEBPPGKSRREOQCKKNAVN 600  
DB 541 RTGGGVOFSHRECDPEPONGRCYCLGRAPKQSCHEBPPGKSRREOQCKKNAVN 600  
QY 601 YVDMQGNLLQWPKYAGVSPRDRCLFCRARGRSEFFKFEAKVIDGLTGGETLALCYRG 660  
DB 601 YVDMQGNLLQWPKYAGVSPRDRCLFCRARGRSEFFKFEAKVIDGLTGGETLALCYRG 660  
QY 661 QCVKAGCDHVDSPRLDKGVCGKNGSKRVSSSLPTNYGNDVYTIIPAGATNDVK 720  
DB 661 QCVKAGCDHVDSPRLDKGVCGKNGSKRVSSSLPTNYGNDVYTIIPAGATNDVK 720  
QY 721 QSHHPVQNDGYLALKTADGQYLLNGNLAISAIQDILVKTILKYSGSATTERLOSF 780  
DB 721 QSHHPVQNDGYLALKTADGQYLLNGNLAISAIQDILVKTILKYSGSATTERLOSF 780  
QY 781 RPLPEPLVQLLTPGEVFPKYKTYFFVNDVDSMOSSKRRATNTIIPILLHAWNLG 840  
DB 781 RPLPEPLVQLLTPGEVFPKYKTYFFVNDVDSMOSSKRRATNTIIPILLHAWNLG 840

QY 841 DWSGCSSTCGAGWQRRVTEGCRDPGGSATCNKALKREDAKPCESQICPL 890  
DB 841 DWSGCSSTCGAGWQRRVTEGCRDPGGSATCNKALKREDAKPCESQICPL 890

# RESULT 2

AA350003 standard; protein, 890 AA.

AA350003;

19-MAR-2001 (first entry)

Human METH2.

Human; METH2; metalloprotease; thrombospondin; angiogenesis inhibition; cancer therapy; benign tumor; ocular angiogenic disease; rheumatoid arthritis; psoriasis; wound healing; endometriosis; vasculogenesis; granulation; hypertrophic scar; nonunion fracture; scleroderma; trachoma; vascular adhesion; myocardial angiogenesis; coronary limb angiogenesis; Osler-Webber syndrome; Crohn's disease; atherosclerosis; birth control.

Homosapiens.

WO200071577-A1.

30-NOV-2000.

25-MAY-2000; 2000WO-US014462.

25-MAY-1999; 99US-00318208.  
20-DEC-1999; 99US-0144882P.  
10-AUG-1999; 99US-0147823P.  
13-AUG-1999; 99US-00373658.  
22-DEC-1999; 99US-0171503P.  
22-FEB-2000; 2000US-0183792P.

(HUMA-) HUMAN GENOME SCI INC.

(SMIR) SMITHKLINE BEECHAM CORP.

(BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

(IRUE) IRUELA-ARISPE L.

(HAST/) HASTINGS G A.

(RUBE/) RUBEN S M.

(JONK/) JONAK Z L.

(TRULL/) TRULLI S H.

(FORN/) FORNWALD J A.

(TERR) TERRRETT J A.

IrueLA-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH; Fornwald JA, Terrett JA;

WPI: 2001-025136/03.

N-PSDB; AAC90058.

METH1 and METH2 polynucleotides and encoded polypeptides, used to inhibit angiogenesis in the treatment of disorders such as cancer, rheumatoid arthritis and psoriasis.

Claim 15; Fig 2; 768pp; English.

The present sequence is human METH2 (ME for metalloprotease and TH for thrombospondin). METH2 can be used for inhibiting angiogenesis in an individual, and for treating cancer, benign tumors, an ocular angiogenic disease, rheumatoid arthritis, psoriasis, delayed wound healing, nonunion fracture, scleroderma, trachoma, vascular adhesions, myocardial angiogenesis, coronary collateral, cerebral collaterals, arteriovenous malformations, ischaemic limb angiogenesis, Osler-Webber syndrome, plaque

CC neovascularisation, telangiectasia, haemophilic joints, angiodysplasia,  
 CC fibromuscular dysplasia, wound granulation, Crohn's disease or  
 CC atherosclerosis. METH2 can also be used in birth control. METH2 can also  
 CC be used in diagnostic methods for the prognosis of cancer

XX Sequence 890 AA;

Query Match 100.0%; Score 4853; DB 4; Length 890;  
 Best local similarity 100.0%; Pred. No. 0;  
 Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFPAAPARWLPFLILLILLILLPLARGAPAPAPAGGASGLVPTPLPSAGELALHLSA 60  
 DB 1 MFPAAPARWLPFLILLILLILLPLARGAPAPAPAGGASGLVPTPLPSAGELALHLSA 60  
 QY 61 FGKGFVRLAPDSDFLAPFKIERLGSGGRATGGERGLRCGFCPSGVNGEPBSLAIVSLC 120  
 DB 61 FGKGFVRLAPDSDFLAPFKIERLGSGGRATGGERGLRCGFCPSGVNGEPBSLAIVSLC 120  
 QY 121 RGLSGSFLLDGESEFTIQPGAGSGLAOPHRLQRMGPAGAPLPRGPEWEVETGEGROER 180  
 DB 121 RGLSGSFLLDGESEFTIQPGAGSGLAOPHRLQRMGPAGAPLPRGPEWEVETGEGROER 180  
 QY 181 GDHODESEESQEEAEAGASEPPPLGATSRTRKFVSEARFVETLLVAASMAAFYADL 240  
 DB 181 GDHODESEESQEEAEAGASEPPPLGATSRTRKFVSEARFVETLLVAASMAAFYADL 240  
 QY 241 QNHILITMSVAARIYKPSIKNSINLMVVKVLLVEDEKMGPEVSDNGGLTLRNFQWQR 300  
 DB 241 QNHILITMSVAARIYKPSIKNSINLMVVKVLLVEDEKMGPEVSDNGGLTLRNFQWQR 300  
 QY 301 QNHILITMSVAARIYKPSIKNSINLMVVKVLLVEDEKMGPEVSDNGGLTLRNFQWQR 300  
 DB 301 QNHILITMSVAARIYKPSIKNSINLMVVKVLLVEDEKMGPEVSDNGGLTLRNFQWQR 300  
 QY 361 TLAHELGLVLSMPHDSKPCRTLFGPMGKHVMAFLFVHLNQLTLPMSPCSAMLTLLDQ 420  
 DB 361 TLAHELGLVLSMPHDSKPCRTLFGPMGKHVMAFLFVHLNQLTLPMSPCSAMLTLLDQ 420  
 QY 421 GHGQCLDAPGAAFLPLTGLPGRNALYQLDQCKQLGPFPRICPNTSAQDCAQLMCHT 480  
 DB 421 GHGQCLDAPGAAFLPLTGLPGRNALYQLDQCKQLGPFPRICPNTSAQDCAQLMCHT 480  
 QY 481 DGAEPLCHTKNGSLPMADGTPCGPGLCSGSCLPSEEVPRPKPVVDGAPMGWGECS 540  
 DB 481 DGAEPLCHTKNGSLPMADGTPCGPGLCSGSCLPSEEVPRPKPVVDGAPMGWGECS 540  
 QY 541 RTGGGVGFHSRECKDPONGGRVCLGRAXOSCHTEPCPDGKSPRQCEKXNAY 600  
 DB 541 RTGGGVGFHSRECKDPONGGRVCLGRAXOSCHTEPCPDGKSPRQCEKXNAY 600  
 QY 601 YTTMDGNLLQWPKYAGVSPDRCKLFCRARGSEFVFAKYIDGTLGSPETLALCVAG 660  
 DB 601 YTTMDGNLLQWPKYAGVSPDRCKLFCRARGSEFVFAKYIDGTLGSPETLALCVAG 660  
 QY 661 QCYVAGCHVVDSPKDKCGVCGGKNSCKXVSGSLTPNNGVNDIVITPAATIDVK 720  
 DB 661 QCYVAGCHVVDSPKDKCGVCGGKNSCKXVSGSLTPNNGVNDIVITPAATIDVK 720  
 QY 721 QRSHPGVNDGNVYALKTADQVLLNGNLAISALBODILVKGTLIXKSGSIATLERLOS 780  
 DB 721 QRSHPGVNDGNVYALKTADQVLLNGNLAISALBODILVKGTLIXKSGSIATLERLOS 780  
 QY 781 RPLPEPLTVOLLTPGVEFPKVKYTFPVNDVDSQSSKERATNTIOLPHAWVIG 840  
 DB 781 RPLPEPLTVOLLTPGVEFPKVKYTFPVNDVDSQSSKERATNTIOLPHAWVIG 840  
 QY 841 DMSECSSTCGAGWQRRIVECRDPGQASATCNKALKEPAKPCESQICPL 890  
 DB 841 DMSECSSTCGAGWQRRIVECRDPGQASATCNKALKEPAKPCESQICPL 890

RESULT 3

ABP96306  
 ID ABP96306 standard; protein: 890 AA.

AC ABP96306;

DT 20-MAY-2003 (first entry)

DE Human ADAMTS8 protein.

KM Humanised baculovirus; cytosstatic; gene therapy; baculovirus; cancer;  
 KM prostate cancer; chromosome 11.

OS Homo sapiens.

PN MO2003016540-A2.

PD 27-FEB-2003.

PF 15-AUG-2002; 2002WO-GB003791.

PR 15-AUG-2001; 2001GB-00019852.

PA (UYYO-) UNIV YORK.

PI Maitland N;

DR WPI; 2003-268336/26.

PT New baculovirus having a modified genome encoding a therapeutic agent,  
 PT useful in the manufacture of a medicament for the treatment of cancer,  
 PT particularly prostate cancer.

PS Claim 24; Page; 34pp; English.

CC The present invention describes a humanised baculovirus (1) which  
 CC comprises a modified baculovirus genome having a nucleic acid molecule  
 CC encoding a therapeutic agent and a polypeptide which functions to target  
 CC the baculovirus to at least one cell type. Also described is a  
 CC pharmaceutical composition comprising (1). (1) has cytostatic activity,  
 CC and can be used in gene therapy. The baculovirus is useful in the  
 CC manufacture of a medicament for the treatment of cancer, particularly  
 CC prostate cancer. The present sequence represents the human disintegrin-  
 CC like and metalloprotease (reprolysin type) with thrombospondin type 1  
 CC motif, 8 (ADAMTS8) protein, which is specified in the exemplification of  
 CC the present invention as angiotensin. N.B. The present sequence is not  
 CC given in the specification but is referred to in Claim 24 as Genbank  
 CC accession number NM\_007037

SQ Sequence 890 AA;

Query Match 99.4%; Score 4822; DB 6; Length 890;

Best local similarity 99.4%; Pred. No. 0;  
 Matches 885; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MFPAAPARWLPFLILLILLILLPLARGAPAPAPAGGASGLVPTPLPSAGELALHLSA 60  
 DB 1 MFPAAPARWLPFLILLILLILLPLARGAPAPAPAGGASGLVPTPLPSAGELALHLSA 60  
 QY 61 FGKGFVRLAPDSDFLAPFKIERLGSGGRATGGERGLRCGFCPSGVNGEPBSLAIVSLC 120  
 DB 61 FGKGFVRLAPDSDFLAPFKIERLGSGGRATGGERGLRCGFCPSGVNGEPBSLAIVSLC 120  
 QY 121 RGLSGSFLLDGESEFTIQPGAGSGLAOPHRLQRMGPAGAPLPRGPEWEVETGEGROER 180  
 DB 121 RGLSGSFLLDGESEFTIQPGAGSGLAOPHRLQRMGPAGAPLPRGPEWEVETGEGROER 180  
 QY 181 GDHODESEESQEEAEAGASEPPPLGATSRTRKFVSEARFVETLLVAASMAAFYADL 240  
 DB 181 GDHODESEESQEEAEAGASEPPPLGATSRTRKFVSEARFVETLLVAASMAAFYADL 240  
 QY 241 QNHILITMSVAARIYKPSIKNSINLMVVKVLLVEDEKMGPEVSDNGGLTLRNFQWQR 300  
 DB 241 QNHILITMSVAARIYKPSIKNSINLMVVKVLLVEDEKMGPEVSDNGGLTLRNFQWQR 300

```

QY 301 FNPSPDRHREHYDTAILLTRONFCGQEGLCDTLGVADIGTICDPNKSQSVIEDGLQAAH 360
DB 301 FNPSPDRHREHYDTAILLTRONFCGQEGLCDTLGVADIGTICDPNKSQSVIEDGLQAAH 360
QY 361 TLNHELGHVLSMHDSPKCTRLFGPMGKHVMAPLFVHLNQLTPMSPCSAMYLTELIDG 420
DB 361 TLNHELGHVLSMHDSPKCTRLFGPMGKHVMAPLFVHLNQLTPMSPCSAMYLTELIDG 420
QY 421 GHGDCILDAFPAALPLPTGLPGEMALYOLDQOCROIFFGDFRHCNPTSADVCALMCHT 480
DB 421 GHGDCILDAFPAALPLPTGLPGEMALYOLDQOCROIFFGDFRHCNPTSADVCALMCHT 480
QY 481 DGAEPILCHTNKSLPMAADGTPCGPHLCSGSLCPBEVEVERPKPVADGMAFWMGESCS 540
DB 481 DGAEPILCHTNKSLPMAADGTPCGPHLCSGSLCPBEVEVERPKPVADGMAFWMGESCS 540
QY 541 RTGCGGVQFSHRECKDPEPONGRKYCLGRBAKYOSCHTEBCPPDGKSPREOQCEKYNAYN 600
DB 541 RTGCGGVQFSHRECKDPEPONGRKYCLGRBAKYOSCHTEBCPPDGKSPREOQCEKYNAYN 600
QY 601 YTTMDGNLLQWVPKTYAGVSPDRCKLFCRARGSEFFVFAKYIDGTLGSPETLALCVRG 660
DB 601 YTTMDGNLLQWVPKTYAGVSPDRCKLFCRARGSEFFVFAKYIDGTLGSPETLALCVRG 660
QY 661 QCVKAGGDHVVDSFRKLDKCGVCGKGNSCRKYSGLTPNNGYNDIVITIPAGATNIDYK 720
DB 661 QCVKAGGDHVVDSFRKLDKCGVCGKGNSCRKYSGLTPNNGYNDIVITIPAGATNIDYK 720
QY 721 QRSHPGVQNDGNVYALKTADGQYLLNGNLAISAEODILVKGTILKXSGSIALTELQSF 780
DB 721 QRSHPGVQNDGNVYALKTADGQYLLNGNLAISAEODILVKGTILKXSGSIALTELQSF 780
QY 781 RPLPEPLTVQLLTVPGSEVFPKXYTFVFNVDVFSMOSSKERATTNIIQPLIHAQWVIG 840
DB 781 RPLPEPLTVQLLTVPGSEVFPKXYTFVFNVDVFSMOSSKERATTNIIQPLIHAQWVIG 840
QY 841 DMSBCSSTCGAGWQRTVTECRDPSSQASATCNKALKPEDAKPCESQJLCP 890
DB 841 DMSBCSSTCGAGWQRTVTECRDPSSQASATCNKALKPEDAKPCESQJLCP 890

RESULT 4
AAB74946
ID AAB74946 standard; protein; 889 AA.
AC AAB74946;
XX 02-JUL-2001 (first entry)
DE Human ADAM type metal protease MDT3 protein SEQ ID NO:20.
XX Human; a disintegrin and metalloprotease type metal protease; MDT3;
XX MDT32; MDT33; ADAM type metal protease; cytosolic; antiarthritic;
XX cancer; arthritis; arthrosis deformans.
XX Homo sapiens.
XX JP2001008687-A.
XX 16-JAN-2001.
XX 25-JUN-1999; 99JP-00180973.
XX 25-JUN-1999; 99JP-00180973.
XX (YAMA) YAMANOUCHI PHARM CO LTD.
XX WPI: 2001-285362/30.
XX N-PSDB; AAF82166.
XX New metal protease and metal protease gene, for use as a drug for
XX cancers, arthritis and arthrosis deformans.
PT

```

```

XX Claim 1; Page 25-27; 31pp; Japanese.
BS The present sequence represents a disintegrin and metalloprotease (ADAM)
XX type metal protease designated MDT3, isolated from human. MDT3 proteins
CC have cytosolic and antiarthritic activities. They can be used as a drug
CC for cancers, arthritis and arthrosis deformans
XX
XX Sequence 889 AA:
Query Match 99.1%; Score 4807.5; DB 4; Length 889;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 884; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

1 MFPAAPARWLPFLILLILLILLPLARGAPAPPAAGGASSELVETRLPGSAGELALHLSA 60
DB 1 MLPAAPARWLP-LLILLILLILLPLARGAPAPPAAGGASSELVETRLPGSAGELALHLSA 59
QY 61 FGKGFVLRPLADDSFLAPDEFKIEELGSGGATGGERLRCGFPGCTVNGPESLAAVSLC 120
DB 61 FGKGFVLRPLADDSFLAPDEFKIEELGSGGATGGERLRCGFPGCTVNGPESLAAVSLC 119
QY 121 RGLSGSFLLDGEEFTIQOGAGSGLAOPHRLQRMWGPAGARPLRGPEWEVETGEGQOER 180
DB 121 RGLSGSFLLDGEEFTIQOGAGSGLAOPHRLQRMWGPAGARPLRGPEWEVETGEGQOER 179
QY 181 GDHQDSEESQEEBAGASPPPLGATSTKAPVESEARFVETLLVADASMAAFYQADL 240
DB 181 GDHQDSEESQEEBAGASPPPLGATSTKAPVESEARFVETLLVADASMAAFYQADL 239
QY 241 QNHILLTMSVAAIKYKPSIKNSINLMVVKYLIVDEKMGVSDNGSLTRNFCNNQRR 300
DB 241 QNHILLTMSVAAIKYKPSIKNSINLMVVKYLIVDEKMGVSDNGSLTRNFCNNQRR 299
QY 301 FNPSPDRHREHYDTAILLTRONFCGQEGLCDTLGVADIGTICDPNKSQSVIEDGLQAAH 360
DB 301 FNPSPDRHREHYDTAILLTRONFCGQEGLCDTLGVADIGTICDPNKSQSVIEDGLQAAH 359
QY 361 TLNHELGHVLSMHDSPKCTRLFGPMGKHVMAPLFVHLNQLTPMSPCSAMYLTELIDG 420
DB 361 TLNHELGHVLSMHDSPKCTRLFGPMGKHVMAPLFVHLNQLTPMSPCSAMYLTELIDG 419
QY 421 GHGDCILDAFPAALPLPTGLPGEMALYOLDQOCROIFFGDFRHCNPTSADVCALMCHT 480
DB 421 GHGDCILDAFPAALPLPTGLPGEMALYOLDQOCROIFFGDFRHCNPTSADVCALMCHT 479
QY 481 DGAEPILCHTNKSLPMAADGTPCGPHLCSGSLCPBEVEVERPKPVADGMAFWMGESCS 540
DB 481 DGAEPILCHTNKSLPMAADGTPCGPHLCSGSLCPBEVEVERPKPVADGMAFWMGESCS 539
QY 541 RTGCGGVQFSHRECKDPEPONGRKYCLGRBAKYOSCHTEBCPPDGKSPREOQCEKYNAYN 600
DB 541 RTGCGGVQFSHRECKDPEPONGRKYCLGRBAKYOSCHTEBCPPDGKSPREOQCEKYNAYN 599
QY 601 YTTMDGNLLQWVPKTYAGVSPDRCKLFCRARGSEFFVFAKYIDGTLGSPETLALCVRG 660
DB 601 YTTMDGNLLQWVPKTYAGVSPDRCKLFCRARGSEFFVFAKYIDGTLGSPETLALCVRG 659
QY 661 QCVKAGGDHVVDSFRKLDKCGVCGKGNSCRKYSGLTPNNGYNDIVITIPAGATNIDYK 720
DB 661 QCVKAGGDHVVDSFRKLDKCGVCGKGNSCRKYSGLTPNNGYNDIVITIPAGATNIDYK 719
QY 721 QRSHPGVQNDGNVYALKTADGQYLLNGNLAISAEODILVKGTILKXSGSIALTELQSF 780
DB 721 QRSHPGVQNDGNVYALKTADGQYLLNGNLAISAEODILVKGTILKXSGSIALTELQSF 779
QY 781 RPLPEPLTVQLLTVPGSEVFPKXYTFVFNVDVFSMOSSKERATTNIIQPLIHAQWVIG 840
DB 781 RPLPEPLTVQLLTVPGSEVFPKXYTFVFNVDVFSMOSSKERATTNIIQPLIHAQWVIG 839
QY 841 DMSBCSSTCGAGWQRTVTECRDPSSQASATCNKALKPEDAKPCESQJLCP 890
DB 841 DMSBCSSTCGAGWQRTVTECRDPSSQASATCNKALKPEDAKPCESQJLCP 889

```

RESULT 5  
 AAB72284  
 ID AAB72284 standard; protein; 905 AA.  
 XX  
 AC AAB72284;  
 XX  
 DT 14-MAY-2001 (first entry)  
 XX  
 DE Murine ADAMTS-8 amino acid sequence.  
 XX  
 KM ADAMTS-N, disintegrin, metalloprotease; thrombospondin type I motif;  
 KM tumour cachexia; inflammation; dermatosparaxis; EDS-VIIC; angiogenesis;  
 KM Ehlers-Danlos syndrome type VIIC; articular cartilage erosion; mouse;  
 KM metastasis; embryogenesis; egg implantation; ADAMTS-8.  
 XX  
 OS Mus musculus.  
 XX  
 PN M0200111074-A2.  
 XX  
 PD 15-FEB-2001.  
 XX  
 PF 03-AUG-2000; 2000MO-US021223.  
 XX  
 PR 06-AUG-1999; 99US-00369364.  
 XX  
 PA (CLEV-) CLEYELAND CLINIC FOUND.  
 PA (APTE-) APTE S. S.  
 PA (HURS-) HURSKAINEN T. L.  
 PA (HIRO-) HIROHATA S.  
 PI  
 PI Apte SS, Hurskainen TL, Hirohata S;  
 DR WPI, 2001-159978/16.  
 DR N-PSDB; AAF63441.  
 XX  
 PT Murine and human 'A Disintegrin-like And Metalloprotease domain with  
 PT Thrombospondin type I motifs' proteins and the nucleic acids encoding  
 PT them, useful for treating e.g. tumors, inflammation and arthritis.  
 XX  
 PS Claim 1; Fig 5; 181pp; English.  
 XX  
 CC This invention relates to murine and human ADAMTS-N (A disintegrin-like  
 CC and metalloprotease domain with thrombospondin type I motifs) proteins,  
 CC designated ADAMTS-5, 6, 7, 8, 9, 10 and R1. Also included in the  
 CC invention are cDNA sequences encoding the proteins, and antibodies  
 CC specific for the proteins. The nucleic acid sequences and proteins may be  
 CC used in the prevention, diagnosis and treatment of diseases associated  
 CC with inappropriate ADAMTS-N expression. Disorders that may be treated  
 CC using the nucleic acids, proteins and antibodies include, for example  
 CC tumour cachexia, inflammation, dermatosparaxis in cattle or Ehlers-Danlos  
 CC syndrome type VIIC (EDS-VIIC) in humans, erosion of articular cartilage  
 CC in arthritic (both inflammatory and non-inflammatory) disease,  
 CC angiogenesis, tumour growth and metastases, and they may also be used for  
 CC controlling embryogenesis and implantation of fertilised eggs. The  
 CC present sequence represents murine ADAMTS-8  
 XX  
 SQ Sequence 905 AA;  
 Query Match 80.8%; Score 3919.5; DB 4; Length 905;  
 Best Local Similarity 79.2%; Pred. No. 3e-269;  
 Matches 717; Conservative 74; Mismatches 99; Indels 15; Gaps 4;  
 QY 1 MFPAAPARWLPFLILLILL-LPLARCAPAPPAAGQASLIVPTRLPAGSLAHLIS 59  
 DB 1 MLRDPITGWPILLILLQLPPEPLVCGAPGPGTGAQASLIVPTRLPAGSLAHLIS 60  
 QY 60 ARGGFVRLAPDDSFILAPFKIERLGGSGATGGERGLRGCFSGTVNGSPESLAAVSL 119  
 DB 61 ARGGFVRLAPDDSFILAPFKIERLGGSSAAAGSPGLRGCFSGTVNGRSLAAMSC 120  
 QY 120 CRGLSGFLLDGEFTIOPGAGGSLAOPHRLQRWPGAGAR-----PLPRPEW 168

DB 121 VAGWSGFLLAGBEFTIQPGAGDSLDPHRLQRWPGGQREDPGLAAAEVFLPQGLEW 180  
 QY 169 EVETEGQROERGDHODESEBEOE--EAEAGSEPPPLGATNRTREFSEARFVETLL 226  
 DB 181 EEWGNGQGOGRSDNEDEKQDEXGLKTEDSKRPFPFGSKTRSRFVSEARFVETLL 240  
 QY 227 VADASMAAFYADLQNHILITLMSVAARIYKPSIKNSINIMVAVKILVEDEKMGPEVSDN 286  
 DB 241 VADASMAAFYGTDLQNHILITLMSVAARIYKPSIRNSVNLVYVAVLIVEKERMGEVSDN 300  
 QY 287 GGLTLRNFQWRRFENQPSDRHPHYDTAILLTRONCGOEGCLDTLGVADIGTCDPNK 346  
 DB 301 GGLTLRNFQWRRFENQPSDRHPHYDTAILLTRONCGOEGCLDTLGVADIGTCDPNK 360  
 QY 347 SCSVTEBEGLOAATLHAEIGLHLSMPHDSKPTRLFQPMGKHVNAPLFVHINQTLFW 406  
 DB 361 SCSVTEBEGLOAATLHAEIGLHLSMPHDSKPTRLFQPMGKHVNAPLFVHINQTLFW 420  
 QY 407 SPCSAMVITTELLDGHGDCILDAPGAALPLPTGLPGRMALYOLDQCRQIFGDPFRICPN 466  
 DB 421 SPCSAMVITTELLDGHGDCILDAPTSVLPPTGLPGRSTLYELDQCRQIFGDPFRICPN 480  
 QY 467 TSADQVCAQLWC-HTDGAEPICHTKNGSLPADGTFCGPGHLCSEGSCLPEEYERPKPV 525  
 DB 491 TSVEDICVQLCARHRDSDEPICHTRKNSLMAADGTPCGFGLCLDGSCLVEDENPKAV 540  
 QY 526 VDGWAPWGPWGECSRTCGGVOFSHRECKDPEPONGRVCYLSRPAKYQCHTEBCPPDG 585  
 DB 541 VDGWGPWRPWPQCSRTCGGIGIFSNECDNPMQNGRFLCIGRRVYKQSNTEBCPPNG 600  
 QY 586 KSFRQCEKYNAYNTMDNLLQWPKYGVSPEDRCXLFCAARGSEKVEAKYID 645  
 DB 601 KSFRQCEKYNAYNTMDNLLQWPKYGVSPEDRCXLFCAARGSEKVEAKYID 660  
 QY 646 GTLCGPETLAIQVAGQCVKAGCDHYVDSPRKLDKCGVCGKNSCRKXVSGSLPTNNGYN 705  
 DB 661 GTLCGPETLAIQVAGQCVKAGCDHYVDSPRKLDKCGVCGKNSCRKXVSGSLPTNNGYN 720  
 QY 706 DIVTPAGATNIDVQSRHPGVQDGNATLAKTADGQVTLNGLNLAISIEDILYKGTIL 765  
 DB 721 DIVTPAGATNIDVQSRHPGVQDGNATLAKTADGQVTLNGLNLAISIEDILYKGTIL 780  
 QY 766 KYSGSIATLERLQSRPPEPLTVQLTVPEVEPPKRYTFEPVNDVDFSMQSSKERAT 825  
 DB 781 KYSGSIATLERLQSRPPEPLTVQLTVPEVEPPKRYTFEPVNDVDFSMQSSKERAT 840  
 QY 826 TIIQPLHQAOWVLGDWSECSSTCGAGQRTVECRDPSQAASATCNKALPREDAKPES 885  
 DB 841 TIIQPLHQAOWVLGDWSECSSTCGAGQRTVECRDPSQAASATCNKALPREDAKPES 900  
 QY 886 QICPL 890  
 DB 901 QICPL 905  
 RESULT 6  
 AAB21251  
 ID AAB21251 standard; protein; 680 AA.  
 XX  
 AC AAB21251;  
 XX  
 DT 23-FEB-2001 (first entry)  
 XX  
 DE Human metalloproteinase ADAMTS-2.  
 XX  
 KM Human; ADAMTS2, metalloproteinase; ADAM;  
 KM a disintegrin and metalloproteinase domain; thrombospondin domain;  
 KM vaccine; neutropenic; neuroprotective; antiParkinsonian; cerebroprotective;  
 KM cytoskeletal; antiarthritic; immunosuppressive; Alzheimer's disease;  
 KM Parkinson's disease; stroke; cancer; arthritis; autoimmune disease;  
 KM brain tumour; brain injury.

OS Homo sapiens.  
 XX MO200053774-A2.  
 XX 14-SEP-2000.  
 XX  
 XX 08-MAR-2000; 2000MO-US006237.  
 XX  
 XX 08-MAR-1999; 99US-00264585.  
 XX  
 XX (NEUR-) NEUROCRINE BIOSCIENCES INC.  
 XX  
 XX Kelnner GS, Clark M, Maki RA;  
 XX WPI; 2000-594326/56.  
 XX N-PSDB; AAA95821.  
 XX  
 XX Polynucleotide encoding novel members of a disintegrin, metalloproteinase  
 PT and thrombospondin domain protein family used to prevent and treat  
 PT Alzheimer's disease, cancer and autoimmune diseases.  
 XX  
 XX Claim 12; Fig 2; 129pp; English.  
 XX  
 XX The present sequence is human metalloproteinase ADAMTS-2. The ADAMTS  
 CC family of proteins is closely related to the ADAM (A Disintegrin and  
 CC Metalloproteinase Domain) family. Members of the ADAMTS family contain a  
 CC thrombospondin domain in addition to the disintegrin and  
 CC metalloproteinase domains found in the ADAMs. ADAMTS polypeptides are  
 CC useful for the manufacture of medicaments for treating conditions  
 CC associated with neuroinflammation and/or neurodegeneration, such as  
 CC Alzheimer's disease, Parkinson's disease and stroke. They are also useful  
 CC for treating conditions associated with cell proliferation, cell  
 CC migration, inflammation and/or angiogenesis, such as cancer, arthritis  
 CC and autoimmune diseases. They can be used to treat patients afflicted  
 CC with an invasive tumour, a brain tumour or brain injury  
 CC  
 XX Sequence 680 AA;  
 XX  
 XX Query Match 76.5%; Score 3713; DB 3; Length 680;  
 XX Best Local Similarity 99.1%; Pred. No. 9,4e-255;  
 XX Matches 674; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 XX  
 XX 211 RTKRFSEARFVETLLVADASMAAFYGLDQNHILTLMSVARTKPSIKNSINLWVK 270  
 XX 1 RTKRFSEARFVETLLVADASMAAFYGLDQNHILTLMSVARTKPSIKNSINLWVK 60  
 XX 271 VLVDEKRGPEVSDNGGLTLNFCNQRRPQSPDRPEHYDRIALLITRONFGGGLC 330  
 XX 61 VLVDEKRGPEVSDNGGLTLNFCNQRRPQSPDRPEHYDRIALLITRONFGGGLC 120  
 XX 331 DTLGVADIGTICDPNKSQVIEDEGLQAAHTLAHGLVSLMPSHDSKPCRTLFGPVGKI 390  
 XX 121 DTLGVADIGTICDPNKSQVIEDEGLQAAHTLAHGLVSLMPSHDSKPCRTLFGPVGKI 180  
 XX 391 HTMAPEFVHLNQTLPMSPCSAMYLTELDDGGHGLDLPAAAPLPFGLPGRALYQLD 450  
 XX 181 HTMAPEFVHLNQTLPMSPCSAMYLTELDDGGHGLDLPAAAPLPFGLPGRALYQLD 240  
 XX 451 QOCROIFGDPFHNCPTNSADQVCAQLMCHTDGABPLCHTKNGSLPMWDGTPCGGHLCS 510  
 XX 241 QOCROIFGDPFHNCPTNSADQVCAQLMCHTDGABPLCHTKNGSLPMWDGTPCGGHLCS 300  
 XX 511 GSCLPHEVERKPYVDGGMAMPWGSCSTCGGVQFSRECKDEBPONGRHYCLGRR 570  
 XX 301 GSCLPHEVERKPYVDGGMAMPWGSCSTCGGVQFSRECKDEBPONGRHYCLGRR 360  
 XX 571 AKYQSCHECPDQSGFRQOCEKXNANYTMDGNLLQVAPKAGVSPDRCKLFCRA 630  
 XX 361 AKYQSCHECPDQSGFRQOCEKXNANYTMDGNLLQVAPKAGVSPDRCKLFCRA 420  
 XX 631 RGRSEKVEAKVITGTLGPEETLACVRCQCYACGCHVYDSRKLDPKCGGKGNSC 690  
 XX 421 RGRSEKVEAKVITGTLGPEETLACVRCQCYACGCHVYDSRKLDPKCGGKGNSC 480

QY 691 RKVSGSLPTNYGVNDIVITIPAGATNIDVKORSHPGVQNDGNVYALATADGOVLLNGLA 750  
 DB 481 RKVSGSLPTNYGVNDIVITIPAGATNIDVKORSHPGVQNDGNVYALATADGOVLLNGLA 540  
 QY 751 ISAIQDILVYGTILKXSGSIATLERLOSRPPLPELTVOQLTYGVEFPKXKTYFFVP 810  
 DB 541 ISAIQDILVYGTILKXSGSIATLERLOSRPPLPELTVOQLTYGVEFPKXKTYFFVP 600  
 QY 811 NDVDFSMQSKERATTNIQPLHQAQWVLGDMSECSSTCGAGQRTVECRDPSGQASAT 870  
 DB 601 NDVDFSMQSKERATTNIQPLHQAQWVLGDMSECSSTCGAGQRTVECRDPSGQASAT 660  
 QY 871 CNKALKPEPDAPCESQLCPL 890  
 DB 661 CNKALKPEPDAPCESQLCPL 680  
 RESULT 7  
 ID AAB21265  
 XX AAB21265 standard; protein; 896 AA.  
 AC AAB21265;  
 DT 23-FEB-2001 (first entry)  
 XX  
 DE Mouse metalloproteinase ADAMTS-1.  
 XX  
 KW Mouse; ADAMTS-1; metalloproteinase; ADAM;  
 KW a disintegrin and metalloproteinase domain; thrombospondin domain;  
 KW vaccine; neurotropic; neuroprotective; antiParkinsonian; cerebroprotective;  
 KW cyostatic; antiarthritic; immunosuppressive; Alzheimer's disease;  
 KW Parkinson's disease; stroke; cancer; arthritis; autoimmune disease;  
 KW brain tumour; brain injury.  
 XX  
 OS Mus musculus.  
 XX  
 XX WO200053774-A2.  
 XX 14-SEP-2000.  
 XX  
 XX 08-MAR-2000; 2000MO-US006237.  
 XX PF  
 XX 08-MAR-1999; 99US-00264585.  
 XX  
 XX (NEUR-) NEUROCRINE BIOSCIENCES INC.  
 XX  
 XX Kelnner GS, Clark M, Maki RA;  
 XX WPI; 2000-594326/56.  
 XX  
 XX Polynucleotide encoding novel members of a disintegrin, metalloproteinase  
 PT and thrombospondin domain protein family used to prevent and treat  
 PT Alzheimer's disease, cancer and autoimmune diseases.  
 XX  
 XX Disclosure; Fig 17; 129pp; English.  
 XX  
 XX The present sequence is mouse metalloproteinase ADAMTS-1. The ADAMTS  
 CC family of proteins is closely related to the ADAM (A Disintegrin and  
 CC Metalloproteinase Domain) family. Members of the ADAMTS family contain a  
 CC thrombospondin domain in addition to the disintegrin and  
 CC metalloproteinase domains found in the ADAMs. ADAMTS polypeptides are  
 CC useful for the manufacture of medicaments for treating conditions  
 CC associated with neuroinflammation and/or neurodegeneration, such as  
 CC Alzheimer's disease, Parkinson's disease and stroke. They are also useful  
 CC for treating conditions associated with cell proliferation, cell  
 CC migration, inflammation and/or angiogenesis, such as cancer, arthritis  
 CC and autoimmune diseases. They can be used to treat patients afflicted  
 CC with an invasive tumour, a brain tumour or brain injury  
 CC  
 XX Sequence 896 AA;  
 XX  
 XX Query Match 48.3%; Score 2344; DB 3; Length 896;





```

QY 14 LILLILLPLARCAPAPPAAGQASBLVPT--RLPG-SAGELALHLSPFGKPVIRLA 70
DB 38 LILASITMLCARGAHGRFTE--EDELIVLPSELRARAGHSTTRILDLDFGQDLHKQ 95
QY 71 PDDSFAPPEFKIERLGGSGRATGGR-----GLRGCFSGTYNGEPESLAASLCRGL 123
DB 96 PDSGLABGFTLQIV---GRKSGSAGCLDPTGDLAHCFYGTYNQPGSAAALSLCEGV 152
QY 124 SSGFLDGEFTLQVQAGGSLAOPHRLQRMGPAPPLPGPFWETGEG---ORGE 179
DB 153 RGAFLQGEFFIOPAP-----GVATERAAPVPEESASAPQPHILRRR 198
QY 180 RGD-----HOEDSEESQEEAEGASEPPPLGA-----TSRTKPFVSEA 219
DB 199 RSGGAKCGVMDDELTPSDSRPESQNRNQMVRDPTPOAGKPSGSGIRKKRFVSS 258
QY 220 RVEVETLVADASMAAFYGADLONELITLMSVAARIYKPSIKXNINLMTYKVIIVEDEK 279
DB 259 RVEVETLVADQSMADFHSGSLKHYLLTFVVAARFYKPSIRNSISLVVVKILVIEBOK 318
QY 280 GPEVDNGGLTLRNFQNNQRRPNQSPDRPHRYDTAILLFRONCQGEGLCDTLGVADIG 339
DB 319 GPEVTSNALTIRNFCNNQKQNSPDRPHRYDTAILLFRQDLCSGH--TCDTLGMADVG 377
QY 340 T1CDPNKSCSVTEDEGLQAAHTLAELGHVLSMFDHDSKPCRTLFGPMGHVMAPLFVH 399
DB 378 TVCDPSRSCSVTEDEGLQAAHTLAELGHVLSMFDHDSKPCRTLFGPMGHVMAPLFVH 437
QY 400 INQTLPMSPCSAMYTELLDGGHGDCLDAPGALPLPTGPGMALYQDDQCRQIFPG 459
DB 438 LDHSGPMSPCSAMYTELLDGGHGDCLDAPGALPLPTGPGMALYQDDQCRQIFPG 495
QY 460 DRRHCPTSAQDVCAQWMC-HTDGAELPCHTKNGSLPMADGTPCGPGLHSCBEGCLPEE 518
DB 496 ESKHCPD--AASTCTTLMCTGSGGLVCCQTKH--FFMADGTSGBEGKMCVSKKCVNKD 551
QY 519 VRRPKPVVDGWAAPMGWGECSRTCGGCVQSHRECKDPEHONGRKYCLGRAPKQOSCHT 578
DB 552 MGHFATPVHSGWGPMPGDDCSRTCGGCVQSHRECKDPEHONGRKYCLGRAPKQOSCHT 611
QY 579 BECPD-DGKSPFEOCCERKNAVNTDMGN--LLQWPKIAGVSPDRCLFPRANGSE 635
DB 612 EDCPPNNGKTFEECCCEAHNEFSKASF-GNEPTEWMPKXAGVSPKDRCLTCEBAGIG 670
QY 636 FVYFAKYIDGLCPETALICVRGQCVKAGCDHVVSFRLDKCGVCGKNGSCRRVSG 695
DB 671 FVYLPKVVDPGRCPSDSTSVCGQCVKAGCDRIISKKKPKDKCGVCGKNGSTCKKMG 730
QY 696 SILTPNYGVNDIVTPAGATNIDVKORSHPVQNDGYALAKTDQGYLLNGMLAISLE 755
DB 731 IYTSRPGHDIIVITPAGATNIEVHRNORSGRNGSFLIRADGYIILNGFTLSLE 790
QY 756 QDILVGTIILKYSISATLERLOSFRPLPELTVQTLVPEGVPPPKKXTFTVPNDVF 815
DB 791 QDLTKGYTLKYSISATLERLOSFRPLPELTVQTLVPEGVPPPKKXTFTVPNDVF 843
QY 816 SMOSSKERATNIIOPPLHAQVYLGWEGSSTCGAGMORTTYEGRPSQAATCNKAL 875
DB 844 ---KKKTESFAVITF--SEWVLEWGESCKTSGSGQRRVYVQCRDINGHPASBECKEY 897
QY 876 KPEDAKPESQQLCP 889
DB 898 KPASTRPCADLP 911

```

```

RESULT 9
AA49501
ID AA49501 standard; protein; 950 AA.
AC AA49501;
XX
XX 10-JAN-2000 (first entry)
XX

```

```

DE Human METH1 protein.
XX Human; METH1; METH2; anti-angiogenic; metalloproteinase thrombospondin;
XX cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
XX angiogenesis inhibitor; abnormal wound healing; inflammation;
XX rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
XX diabetic retinopathy; macula degeneration; hemangioma; detection;
XX arterial-venous malformation; immune deficiency.
XX Homo sapiens.
XX WO937660-A1.
XX 29-JUL-1999.
XX 22-JAN-1999; 99MO-US001313.
XX 23-JAN-1998; 98US-0072288P.
XX 28-AUG-1998; 98US-0098539P.
XX (IRUE/) IRUELA-ARISPE L.
XX (HAST/) HASTINGS G A.
XX (RUBE/) RUBEN S M.
XX IrueLA-Arispe L, Hastings GA, Ruben SM;
XX WPI, 1999-590684/50.
XX DR N-PEDB; AA232000.
XX New isolated metalloproteinase thrombospondin polypeptides, useful for
XX treating hyperproliferative disorders, cancers or autoimmune disorders.
XX Claim 10; Fig 1; 457bp; English.
XX AA232000 and AA232001 encode, and AA49501 and AA49502 represent, human
XX metalloproteinase thrombospondin (METH) proteins METH1 and METH2
XX respectively. METH1 and METH2 have been found to be potent inhibitors of
XX angiogenesis both in vitro and in vivo. They can be used for treating
XX cancer and other disorders related to angiogenesis including abnormal
XX wound healing, inflammation, rheumatoid arthritis, psoriasis, endometrial
XX bleeding disorders, diabetic retinopathy, some forms of macula
XX degeneration, haemangiomas, and arterial-venous malformations. They may
XX be useful in treating deficiencies or disorders of the immune system, by
XX activating or inhibiting the proliferation, differentiation, or
XX mobilization (chemotaxis) of immune cells. The etiology of these immune
XX deficiencies or disorders may be genetic, somatic, such as cancer or some
XX autoimmune disorders, acquired (e.g. by chemotherapy or toxins), or
XX infectious. They can also be used to treat inflammatory conditions, both
XX chronic and acute conditions. The products can also be used for detection
XX and diagnosis. AA232002 to AA232080, and AA49503 to AA49511 represent
XX sequences given in the exemplification of the present invention
XX
XX Sequence 950 AA;
XX
XX Query Match 48.0%; Score 2328.5; DA 2; Length 950;
XX Best Local Similarity 49.7%; Pred. No. 2,7e-156;
XX Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;
XX
XX 11 LPTLILLILLPLARCAPAPPAAGQASBLVPT--TRLPASAGELALHLSPFGKPVIR 68
XX 16 VPTLILLALALAVS-DALGRPSE--EDELIVPELERAPGH-GTTRILHAFDQDLJE 71
XX 69 LAPDSEFAPPEFKIERLGGSGRATG-----ERGLGCFSGTYNGEPESLAASLCRGL 123
XX 72 LRPDSFLABGFTLQV---GRKSGSAGCLDPTGDLAHCFYGTYNQPGSAAALSLCEGV 128
QY 124 SSGFLDGEFTLQVQAGGSLAOPHRLQRMGPAPPLPGPFWETGEG---ORGE 172
DB 129 RGAFLYLLGEAVYFIOPAPAS-----ERLAVAPAGEXPPAPLOPHILRRRQDDVGTQCV 183
QY 173 GEGQROERGDHOEDSEESQEEAEGASEPP-----PLGATS--TKKPFVSEARFV 222
DB 184 VDDEPRPTGAETEDDEDEGTGEDEGPGWSPQDPALGVGVGPTGTSTIKKGFVSSHRV 243

```

QY 223 ETLVADASMAAFYADLQNHILTLMSVAARIYKHPISIKNSINIMVVKYLVEDEKMGPE 282  
 DB 244 ETLVADASMAAFYADLQNHILTLMSVAARIYKHPISIKNSINIMVVKYLVEDEKMGPE 303  
 QY 283 VSDNGGLTLNFCNMQRFPNPSDRHPEHYDTALLTRQNFQCGEGCLDPLGVADIGTIC 342  
 DB 304 VTSNAALTLNFCNMQRFPNPSDRHPEHYDTALLTRQNFQCGEGCLDPLGVADIGTIC 362  
 QY 343 DPNKSCSVIEDBGLQAAHTLAHELGHVLSMPHDSKPCRTLFGPMGKHVYADLFVHLNQ 402  
 DB 363 DPNKSCSVIEDBGLQAAHTLAHELGHVLSMPHDSKPCRTLFGPMGKHVYADLFVHLNQ 422  
 QY 403 TLPWSPGSAWYLTLLDGGHDDCLLDAGALPLPTGLPGPMALYQDDQCRQIFGPDPR 462  
 DB 423 SQPWSFGSAWYLTLLDGGHDDCLLDAGALPLPTGLPGPMALYQDDQCRQIFGPDPR 480  
 QY 463 HCPNTSAQDVCAQLWC-HTDGAEPICHTKNGSLPMWADGTGCGPHLCSBGSCLPEEVEER 521  
 DB 481 HCPD--AASCTSLMCTGTSGVLVCOYTKH--FPMADGTGCGEGKMCINGKCVNKTDRKH 536  
 QY 522 PKPVVDGMAFPMGMBGCSRTCGGGVOPSHRECKDEPBGNGRYCGRAXKYOSCHTEEC 581  
 DB 537 FDTFPHSGWGMGMBGCSRTCGGGVOPSHRECKDEPBGNGRYCGRAXKYOSCHTEEC 596  
 QY 562 PP-DGKSFRFEOQCEKRYNAYTDM-DGNLQWPKYAGVSPRDRCKLFCRARGSBPKYF 639  
 DB 597 PDNNGKTFREOQCEKRYNAYTDM-DGNLQWPKYAGVSPRDRCKLFCRARGSBPKYF 656  
 QY 640 EAKYIDGLGPELTALCVRGQCVKAGCDHVDSBRLDKQVCGGKNSCRVSGSLRP 699  
 DB 657 QPKVVDGTPGSPDSTSVCGQCVKAGCDHVDSBRLDKQVCGGKNSCRVSGSLRP 716  
 QY 700 TNYGNDIVTIPAGATINDVQRSHPVQNDGNVLTALKTADGQYLLNGMLAISALBODIL 759  
 DB 717 AKPGYHDIITPTGATNIEVKQRNQRSGNNGSFLAIKADGTYLLNGDYTLSTLBODIL 776  
 QY 760 VKGTLTKXSGSLATLERQSPRLPEPLTVOLLTVPEBEVFPKYKTYFFPNVDVDSMGS 819  
 DB 777 YKGVNLTYSGSALERTSRFSPLEKPEPLTVOLLTVPEBEVFPKYKTYFFPNVDVDSMGS 825  
 QY 820 SKERATNIIOLPLLAQWVLDWMBSCSSTCGAQNRRTVECRDPSGQASATCNKALKP 879  
 DB 826 KKKKSPNAL-PTFSA-WVIEWMBSCSCELQGRSLVECRDINGPASECAKRYKAPAS 883  
 QY 880 AKPCESQICP 889  
 DB 884 TRPCADHPCP 893

RESULT 10  
 AAB73549  
 ID AAB73549 standard, protein, 950 AA.

XX AAB73549;  
 AC AAB73549;  
 DT 07-AUG-2001 (first entry)  
 XX Human ADAM-type metalloprotease MDTs4, SEQ ID NO:4.  
 DE Human: MDTs4; ADAM-type metalloprotease; drug screening;  
 KW A Disintegrin And Metalloprotease; cancer; arthritis.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 PN JP2001017183-A.  
 PD 23-JAN-2001.  
 XX 09-UTL-1999; 99UTP-00196584.  
 XX 09-UTL-1999; 99UTP-00196584.  
 PR 09-UTL-1999; 99UTP-00196584.

PA (YAMA) YAMANOCHI PHARM CO LTD.  
 XX WPI; 2001-275950/29.  
 DR N-PSDB; AAH20224.  
 XX A new metal protease and its preparation for use as an anti-cancer and  
 PT anti-arthritis therapeutic.  
 XX Claim 1; Page 12-14; 22pp; Japanese.  
 CC The invention relates to the novel human ADAM (A Disintegrin And  
 CC Metalloprotease)-type metalloproteases MDTs4 (AAB73549) and MDTs5  
 CC (AAB73550). The metalloproteases can be used for the treatment of cancers  
 CC and arthritis. The invention also relates to the genes encoding MDTs4 and  
 CC MDTs5, vectors and host cells containing the MDTs4 or MDTs5 genes, the  
 CC recombinant production of MDTs4 and MDTs5, and antibody specific for  
 CC MDTs4 or MDTs5, and methods of screening for compounds which modulate the  
 CC activity of MDTs4 and/or MDTs5. The present sequence represents human  
 CC MDTs4  
 CC  
 SQ Sequence 950 AA;  
 Query Match 48.0%; Score 2328.5; DB 4; Length 950;  
 Best Local Similarity 49.7%; Pred. No. 2.7e-156;  
 Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;  
 11 LPLULLLULLLPLARCAPAPAGQASELVP--TRLPASAGELALHLISAFKGFVLR 68  
 16 VPTLLLLAALLAVS-DALGRPSE--EDELVVELEBPARGH-GTTRLHLHFDQDLDE 71  
 69 LAPDPSFLAPFKERLKGSGRATG----ERGLAGCFPGSGTVNBPSPSLAVALSCRG 123  
 72 LRPDSFLAPFTLQNV--GRKSGETPLPBTDLNHCRTSGVINGDPSSAALLSCRG 126  
 124 SGSFLDGEETFLQPCGAGSLAOPHLQRMGAPARPL-----RGPEWV---ET 172  
 129 RGAFTYLLGEAVFIQPLPAAS-----ERLATVADGKXPAPLQFHLRRRQGVGTCGV 183  
 173 GEGROGERQHDSE 222  
 184 VDDEPRFTGAETDE 243  
 223 ETLVADASMAAFYADLQNHILTLMSVAARIYKHPISIKNSINIMVVKYLVEDEKMGPE 282  
 DB 244 ETLVADASMAAFYADLQNHILTLMSVAARIYKHPISIKNSINIMVVKYLVEDEKMGPE 303  
 QY 283 VSDNGGLTLNFCNMQRFPNPSDRHPEHYDTALLTRQNFQCGEGCLDPLGVADIGTIC 342  
 DB 304 VTSNAALTLNFCNMQRFPNPSDRHPEHYDTALLTRQNFQCGEGCLDPLGVADIGTIC 362  
 QY 343 DPNKSCSVIEDBGLQAAHTLAHELGHVLSMPHDSKPCRTLFGPMGKHVYADLFVHLNQ 402  
 DB 363 DPNKSCSVIEDBGLQAAHTLAHELGHVLSMPHDSKPCRTLFGPMGKHVYADLFVHLNQ 422  
 QY 403 TLPWSPGSAWYLTLLDGGHDDCLLDAGALPLPTGLPGPMALYQDDQCRQIFGPDPR 462  
 DB 423 SQPWSFGSAWYLTLLDGGHDDCLLDAGALPLPTGLPGPMALYQDDQCRQIFGPDPR 480  
 QY 463 HCPNTSAQDVCAQLWC-HTDGAEPICHTKNGSLPMWADGTGCGPHLCSBGSCLPEEVEER 521  
 DB 481 HCPD--AASCTSLMCTGTSGVLVCOYTKH--FPMADGTGCGEGKMCINGKCVNKTDRKH 536  
 QY 522 PKPVVDGMAFPMGMBGCSRTCGGGVOPSHRECKDEPBGNGRYCGRAXKYOSCHTEEC 581  
 DB 537 FDTFPHSGWGMGMBGCSRTCGGGVOPSHRECKDEPBGNGRYCGRAXKYOSCHTEEC 596  
 QY 562 PP-DGKSFRFEOQCEKRYNAYTDM-DGNLQWPKYAGVSPRDRCKLFCRARGSBPKYF 639  
 DB 597 PDNNGKTFREOQCEKRYNAYTDM-DGNLQWPKYAGVSPRDRCKLFCRARGSBPKYF 656  
 QY 640 EAKYIDGLGPELTALCVRGQCVKAGCDHVDSBRLDKQVCGGKNSCRVSGSLRP 699  
 DB 657 QPKVVDGTPGSPDSTSVCGQCVKAGCDHVDSBRLDKQVCGGKNSCRVSGSLRP 716



Db 717 AKPGYHDIITPTGATNIEVKQRNQRSGNSFLAIKADGTYIINGDYTLSTLEQDIM 776  
 QY 760 VKGTLIKYSGSIATLERLOSFRPLPEPLTVQLLTVGEVFPKVKYTFEPNDVDFSMOS 819  
 Db 777 YKGVVLRYSGSAALEIRSRSPLEKEPLTIQVLTV-GNALRPKIKYTFV----- 825  
 QY 820 SKERATNTIQLPLHQAQVLDWSECSSTCGAGMQRRTVCRRPSCQASATCNKALKPED 879  
 Db 826 KKKKESFNAL-PTPSA-WVLEWGECSKCELGMRRLVBCRDINGQAPASECAKEVKEPAS 883  
 QY 880 AKPCESQLCP 889  
 Db 884 TRPCADHPCP 893

RESULT 12  
 AAW80285 standard; protein; 967 AA.  
 ID AAW80285 standard; protein; 967 AA.  
 AC AAW80285;  
 DT 19-JAN-1999 (first entry)  
 DE Human integrin ligand polypeptide ITGL-TSP.  
 KW ITGL-TSP; integrin ligand; angiogenic disease; cancer; atherosclerosis;  
 chronic inflammatory disorder; rheumatoid arthritis; tissue remodeling;  
 macular degeneration; diabetic retinopathy; Alzheimer's disease; human;  
 restenosis.  
 OS Homo sapiens.  
 PN EP874050-A2.  
 PD 28-OCT-1998.  
 PF 27-JAN-1998; 98EP-00300575.  
 PR 24-APR-1997; 97US-00845496.  
 PA (SMIX) SMITHKLINE BEECHAM CORP.  
 PA (SMIX) SMITHKLINE BEECHAM PLC.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Jonak ZL, Trullis SH, Fromwald JA, Hastings GA, Terrett JA;  
 DR WPI; 1998-544643/47.  
 DR N-PEDB; AAV66508.  
 PT DNA encoding integrin ligand polypeptide ITGL-TSP - used to treat  
 angiogenic diseases, restenosis, Alzheimer's disease and in tissue  
 remodeling.  
 PS Claim 11, Page 6-9; 24pp; English.  
 CC This represents a human integrin ligand polypeptide ITGL-TSP. ITGL-TSP  
 CC polypeptides can be used in the treatment of angiogenic diseases such as  
 CC cancer, cancer metastasis, chronic inflammatory disorders, rheumatoid  
 CC arthritis, atherosclerosis, macular degeneration or diabetic retinopathy,  
 CC restenosis, Alzheimer's disease and tissue remodeling. They can be used  
 CC to treat a subject in need of enhanced activity or expression of the ITGL  
 CC -TSP polypeptide  
 CC  
 SQ Sequence 967 AA;

Query Match 48.0%; Score 2328.5; DB 2; Length 967;  
 Best Local Similarity 49.7%; Pred. No. 2.8e-156;  
 Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;

QY 11 LPFLLILLPLLRGAPAPAGQASELYVP--TRLSGAGELALHLSAFQKGFYLR 68  
 Db 33 VFTLLLAALAAVS-DALGRPSE--EDSELVPELEERAPGH-GTTRLRLHAFDQDLDE 88

QY 69 LAPDSEFLAEPEKIERLGSGGRATG-----ERGLNGCFSSGTVNGEPESLAASLCRGL 123  
 Db 89 LRPFSSFLAAGFTLQNV--GRKSGSEPLPEFDLHCHCYSGSTVNGDPSSAALSLICEYV 145  
 QY 124 SGEFLDGESEFTLOPQAGAGSLAQPHRLQRMWGPAGARPLP-----RGSEMEV-----ET 172  
 Db 146 RGAFTYLIGEVYFIQPLPAAS-----ERLATAPAGEPRPPLQFHLRLRRQDVGCTGV 200  
 QY 173 GSGQRQSRGHQDSESESOEBEAGSEPP-----PLGATS-RTKAFVSARFV 222  
 Db 201 VDDEPRPTGAKTEDEDEGEDEBGPQWSPQDPALQVGPFTGTSIRKKRFVSHRYV 260  
 QY 223 ETLVADASMAAFYADLQNHILTLMSVAARLYKPSIKNSINLAWVKLYIYDEKMGPE 282  
 Db 261 ETMLVADQSMAEFHSGGLKXLYLTLFSVAARLYKPSIKNSINSLVYVKILVTHDEQKPE 320  
 QY 283 VSDNGSLTLRNPQNRFRNPQPSDRHPHYDTAILITRONFCGQEGLCDTLGVDIGTYC 342  
 Db 321 VTSNAALTLENFCNQKQHNPPESDRDAEHYDTAILFTRODLGSGQ-TCDTLGMADVGTVC 379  
 QY 343 DPNKSCSVIEDEGLQAAHTLAHELGHVLSMPHDSKPCRTLPGPMKHHVMAPLFVHLNQ 402  
 Db 380 DSRKSCSVIEDGLQAAFTTAHELGHVFNPHDDAQCSLNGVANDSMMSMLSLNDH 439  
 QY 403 TLPWSPCSAMYLTLLDGHGHDCLDAPALALPLPTGLPGRMALYQLDQCGQIFQDPFR 462  
 Db 440 SQPSPCSAYVITSLFNDNGHECTMDKPNPQLPGLDLPG--TSYDANRQCQFTFGEDESK 497  
 QY 463 HCPVTSADVCQALWC-HTDGAEPILCHTNKGLPAMDGTPCGPCHLCSGSGCLPEBEYVR 521  
 Db 498 HCPD--AASCTSLMCTGTSGLVVCQTKH--FPMAAGTSCBEGKMCITGKVNKTRGH 553  
 QY 522 PKPVVDGMAFWGPMGECSTTCGGVQVFSHRECKDPEPONGRYCLGRPAKYQSCHTEEC 581  
 Db 554 FDTFFHSGMWGMGPMGDCSTTCGGVQVYTRBEDCNFVPNNGKYEGBKVRYSNLEBDC 613  
 QY 582 PP-DGKSFREOCEKKNAYVITDM--DGNLLQWPKYAGVSPDRCKLFCRARGSEPKYF 639  
 Db 614 PDNNGKTFREOCEAHNEFSKASFGSGPRAVENIPKTAGVSPDRCKLQAKGIGTFVYL 673  
 QY 674 QPKVNDGTPCSPDSTGVCVQGCVKACDRIIDSKKFKPCVCGGNGSTCKKIGSAYTS 733  
 QY 700 TNYGNDIVITIPGATNIDVQRSHPGVQNDGNYALKTADQOYLINGLALSALEQDIL 759  
 Db 734 AKPGYHDIITPTGATNIEVKQRNQRSGNSFLAIKADGTYIINGDYTLSTLEQDIM 793  
 QY 760 VKGTLIKYSGSIATLERLOSFRPLPEPLTVQLLTVGEVFPKVKYTFEPNDVDFSMOS 819  
 Db 794 YKGVVLRYSGSAALEIRSRSPLEKEPLTIQVLTV-GNALRPKIKYTFV----- 842  
 QY 820 SKERATNTIQLPLHQAQVLDWSECSSTCGAGMQRRTVCRRPSCQASATCNKALKPED 879  
 Db 843 KKKKESFNAL-PTPSA-WVLEWGECSKCELGMRRLVBCRDINGQAPASECAKEVKEPAS 900  
 QY 880 AKPCESQLCP 889  
 Db 901 TRPCADHPCP 910

RESULT 13  
 AAB50011 standard; protein; 968 AA.  
 ID AAB50011  
 AC AAB50011;  
 DT 19-MAR-2001 (first entry)  
 DE Protein; SEQ ID 125.  
 KW Human, MENHL, metalloprotease; thrombospondin; angiogenesis inhibition;  
 cancer therapy; benign tumour; ocular angiogenic disease;

KW rheumatoid arthritis; psoriasis; wound healing; endometriosis;  
 KW vasculogenesis; granuloma; hypertrophic scar; nonunion fracture;  
 KW scleroderma; trachoma; vascular adhesion; myocardial angiogenesis;  
 KW coronary collateral; cerebral collateral; arteriovenous malformation;  
 KW ischaemic limb angiogenesis; Oler-Webber syndrome;  
 KW plaque neovascularisation; telangiectasia; haemophilic joint;  
 KW angiodysplasia; fibromuscular dysplasia; wound granulation;  
 KW Cronh's disease; atherosclerosis; birth control.  
 OS Homo sapiens.  
 XX WO200071577-A1.  
 XX 30-NOV-2000.  
 XX 25-MAY-2000; 2000WO-US014462.  
 XX 25-MAY-1999; 99US-00318208.  
 XX 20-JUL-1999; 99US-0144882P.  
 XX 10-AUG-1999; 99US-0147823P.  
 XX 13-AUG-1999; 99US-00373658.  
 XX 22-DEC-1999; 99US-0171503P.  
 XX 22-FEB-2000; 2000US-0183792P.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX (SMIK) SMITHKLINE BEECHAM CORP.  
 XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.  
 XX (IRUE/) IRUELA-ARISPE L.  
 XX (HAST/) HASTINGS G A.  
 XX (RUBEN) RUBEN S M.  
 XX (JONAK) JONAK Z L.  
 XX (TRULI) TRULI S H.  
 XX (FORN/) FORNWALD J A.  
 XX (TERR/) TERRRETT J A.  
 XX IrueLA-Arispe L, Hastings GA, Ruben SM, Jonak ZL, Trulli SH;  
 XX Fornwald JA, Terrrett JA.  
 XX WPI, 2001-025136/03.  
 XX METH1 and METH2 polynucleotides and encoded polypeptides, used to inhibit  
 XX angiogenesis in the treatment of disorders such as cancer, rheumatoid  
 XX arthritis and psoriasis.  
 XX Claim 15; Page 759-763; 768pp; English.  
 CC The present invention relates to human METH1 and METH2 (ME for  
 CC metalloproteinase and TH for thrombospondin; see AAB50002 and AAB50003).  
 CC METH can be used for inhibiting angiogenesis in an individual, and for  
 CC treating cancer, benign tumours, an ocular angiogenic disease, rheumatoid  
 CC arthritis, psoriasis, delayed wound healing, endometriosis,  
 CC vasculogenesis, granuloma, hypertrophic scars, nonunion fractures,  
 CC scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,  
 CC coronary collaterals, cerebral collaterals, arteriovenous malformations,  
 CC ischaemic limb angiogenesis, Oler-Webber syndrome, plaque  
 CC neovascularisation, telangiectasia, haemophilic joints, angiodysplasia,  
 CC fibromuscular dysplasia, wound granulation, Cronh's disease or  
 CC atherosclerosis. METH can also be used in birth control. METH can also be  
 CC used in diagnostic methods for the prognosis of cancer. The present  
 CC sequence is a protein isolated in the present invention  
 XX Sequence 968 AA;  
 SQ  
 Query Match 48.0%; Score 2328.5; DB 4; Length 968;  
 Best Local Similarity 49.7%; Pred. No. 2.8e-156;  
 Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;  
 QY 11 LPELLILLLILLLPARGAPAPPAAGQASGLVVP--TLPSAGELALHLSAFGKGFVLR 68  
 DB 34 VPTLLILALALAVS-DALGRPS--EDBELVPELEBAPRH-GTTLRLHAFQGDLE 89  
 QY 69 LAPDSFLAPERIERLGGSGPATG-----ERGLRGCFPGTNGEPESLAASLCRGL 123

DB 90 LRPDSFLAPGFTLQNV--GRKSGSETPLPETDLACHFYSGTNGDPSSAALLSLCEGV 146  
 QY 124 SGLSFLDGEETFIPOGAGSLAPPHLQGWGAPARPLP-----RGPEWEV---ET 172  
 DB 147 RGAIVYLGEAVFIQPLPAS-----ELTAPAAAGEPPAPLQHLRRRRQGVGTCGV 201  
 QY 173 GEGORGERGDHQBESBESQEBEAKASEPP-----PPGANS-RTKRFVSAKRV 222  
 DB 202 VDDEPRPTGAETDEDETEGEGDEGQWSPDPALQGVQPTGTSIRKKRFVSSHRVY 261  
 QY 223 ETLIVADASMAAFYADLQNHILTLMSVAARIYKFSIKNSINTAVYKYLIVEDKMGPE 282  
 DB 262 ETMLVADQSMAPFGSGSLKHYLLTLPVAVARLKHPSINSVSLVYVKLLVHDEKQGE 321  
 QY 283 VSDNGGLTLRNFPCWQRRFNQPSGDRPEHYDTNLLTRQNFQGBELCTLLGVADIGTIC 342  
 DB 322 VTSNAALTLRNFQWQKQHPSPDRDAEHYDTLILTRQDLQSQ--TCDTLQVADGVIC 380  
 QY 343 DPNKSGVIEDGELQAAHTLAHELGHVLSMHDSPKCTELFGPMKHHVMAPLFYHLNQ 402  
 DB 381 DPNKSGVIEDGELQAAHTLAHELGHVFMNPHDQACASLNGVNDSHMMSMLNDH 440  
 QY 403 TLPWSPGSAHYLTELDDGGHDDLDPAGALPLPGLFGKRALYQLDQCCROIPEPDR 462  
 DB 441 SQPWSFCSAYMTSFLDNGHGECIMDKPQNPQLPDDLPG--TSYDANQCCPTFEBSK 498  
 QY 463 HCPNTSADVQACLMC-HTDGAELPCHTKNGSLPMADGPECGPHLCEBSCCLPEBEVER 521  
 DB 499 HCPD--AASTCSITLWCTGSGVIVCOTKH--FPMDGTSCEGKMKCINGKCNKTRDRH 554  
 QY 522 PKPVVDGMAWPGWGCSTCCGGVQFSTRCKDPEPQNGARYCLGRAPKQSCITEEC 581  
 DB 555 FDTPFHSGWGMGPGWGCSTCCGGVQYTRKEDNVPKNGKRYCKGKRYRSCNLEDC 614  
 QY 582 PP-DKSFREQCEKYNAYNYTDM-DGNLQWPKTAGVSPDRCLFCRARGSEFKYF 639  
 DB 615 PDNNKTFPREQCEAHNEPFAKSPGSPAYEMLPKTAGVSPDRCLIQAGIGVFFVL 674  
 QY 640 EAKYIDGTLGPEFLATCYRGCCYKAGCGHNVDSPKLDKCGYCGGKNSCRVSGSLTP 699  
 DB 675 QPKYVDGTPGSPDSTSCVGGQCYKAGCDRIIDSKKFPKCGVGGNGSTCKKISSVTS 734  
 QY 700 TNYGNDIVTIPAGATNIDVQKSHSPGVNDGNYLAKLTADGGYLLNGMLASALIEQDIL 759  
 DB 735 AKPEGHITITPITGATNIEVKQNRNGSRNNGSFLAKADGTYLNGDYTLSTLEQDIM 794  
 QY 760 VKGTYLKYSGLATLELQSPRLPEPLTYQLITVGEVPPPKYKTFPVNDVDSMGS 819  
 DB 795 YKGVYLRYSGLSALERIRFSPKEPLTIOVLV--GNALRPKITYFV----- 843  
 QY 820 SKERATNIIPLHQAQVLDWSECSSTGACQMRRTVECRDPGQASATNCKALKPED 879  
 DB 844 KKKESFPAL-PTPSA-VWIEWGECKSCGELDMQRLVACRDINGQPASECAKVPAS 901  
 QY 880 AKPGESQICP 889  
 DB 902 TRPCADHPCP 911  
 RESULT 14  
 AA04142  
 ID AA04142 standard; protein; 967 AA.  
 XX AA04142;  
 AC AA04142;  
 XX 15-JUN-1999 (first entry)  
 DT 15-JUN-1999 (first entry)  
 XX  
 DE Human Tango-71 protein.  
 XX Human Tango-71 protein.  
 KW Human; Tango-71; Tango-73; Tango-74; Tango-76; Tango-83; diagnosis;  
 KW detection.  
 XX Homo sapiens.

**Pt** New TANGO polypeptides and nucleic acids encoding them - useful as diagnostic agents and for treating disorders caused by aberrant expression of TANGO.

**Cc** Claim 8; Fig 1; 84pp; English.

The present sequence represents human Tango-71. Tango polypeptides are useful for identifying compounds which bind the polypeptide via direct binding, competition binding assays or Tango-71. -73, -74, 76 or -83-mediated signal transduction. Tango polypeptides are also useful for identifying modulating compounds by determining effect on Tango activity. CC Tango polypeptides and nucleic acids are useful for diagnosing diseases related to aberrant expression of Tango, and Tango polypeptides are useful for raising antibodies which can be used in diagnostic assays for detection of Tango, and also for generating anti-idiotypic antibodies for prevention and protection

**Sq** Sequence 967 AA;

**Xx**

Query Match 48.0%; Score 2327.5; DB 2; Length 967;  
Beet Local Similarity 49.7%; Pred. No. 3.3e-156;  
Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22

11 LPELLILLILPLKAGPARPAPGQASELVP--TRPSAGEIALHLSPKGFTLR 68  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
33 VPTLIILAAILAVS-DALRRSE--EDEELVVELERAPGH-GTIRLIAHFDQLLE 88  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
69 LAPDSSFLAPEKIETGLSGSRATGS----EKLRCGFSGTVNGEPRLSAVSICRG 123  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
89 LPDDSFLLAPGFILOVV---GRKSGETLPFETDLACHFYSGIVGDPSAAALSICEVY 145  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
124 SGSEFLIDGEEFTTOPOGAGSLAQPRRIORWGPAGARPLP-----RGEWEV---ET 172  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
146 RGAFFYLGGAYFIQLPAAS-----ERLAITAAGERKPAPLOFHLLRRNRQGVCISGV 200  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
173 GGQQRCEKDHDSDSEESQEBAEGASEPP-----PRLGANS-RTKRPVSARAV 222  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
201 VDDEPKTAKEIETEDEDTEGEDDEBPQWSPODPALQVGQPGTGTSIKKRKRVSSHRY 260  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
223 ETLVLVDAMAATVAGADLNHTLSVAARIKYKHSINKSIIMLVVKVLYVEDEMGPE 282  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
261 ETMLVLVDQSVAEFHGSKFKHYLTLLTSVAARKLKHSINRSVSLVYKKILVIHDEQKGE 320  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
283 VSNDGELLTRNFCNMQRFRFNOPSDRHREPHYTVALITRONFGGOBELCTTLGVADI GTTC 342  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
321 VTSNALLTRNFENCMQKHNPSPDRDAEHYDTAILETTRODLCGSO-TCDILGANADVGYC 379  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
343 DPKKCSGVIEDGEOAHNTLAHNLGNVLSMPHDSRCTLFPMGKNHVMADELFWHLNQ 402  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
380 DPERSCGSVIEDDGIOAFTTAHELGIYFMNHDDAACASLVNGVNDSHMAMLSNI DH 439  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
403 TLWPSPCANMYTTELIDGGHCCLLDAPGALPLFTGLPERMALYOLDQCROIFEDDFR 462  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
440 SQMWPCSAAMYITSFLNDNGHECCIMDKPNPIQLPGDLPg--TSYANRCQQCFTFGDSK 497  
| : | : | : | : | : | : | : | : | : | : | : | : | :  
463 HCNVTASQDVCAQLMC-HTDGAERPLCHTKXGSLPWMDGTCPGQHLCSEGSCUPREEVER 521

Db	498	HCPD--AAATGCTLMCTGISTGGVLCVOTKH--FPMADGTCGSEBKCWICNGCKVAKTRKH	553
QY	522	PKFVVDGMAAPMGWGECSRTCCGGVOPSHRECKDPEPONGRCICLRRAAYOSCHTEEC	581
Db	554	FDTPFHSGWGMGPMGJCSRTCCGGVGYTWRECNPAVKGKGYCEGRVAYRSCNLEDC	613
QY	582	PP-DGKGFREBOCEKRYAAYVTDN-DGNLLQWVYKXAVGSPERDRCKLFCRARGSEFKVF	639
Db	614	PDNNGKTPREQCAAHNEFSAKSGSGSPAVEMWIKYAGVSPKORCKILCQAKGIQYFVL	673
QY	640	EAKYIDGTCGPETLAI CVRGOQCVKAGCDHVDSPRKIDKCGVCGKGNCSKRYSGSLTP	699
Db	674	QPKVVDGPHCPSDSTSCVGOQCVKAGCDRIIDSKKFFDKCGVCGGSGSTCKKISGVTS	733
QY	700	TNVCYNDIVTIPAAATNIDVKORSHPGVQNDGNATLKTADGOYLNLNGLAISAEODIL	759
Db	734	AKPGHDIITPTGATNILEVQRNQRGSRNNGSLTAIAAGTYILINDGYLTSLTEQDIM	793
QY	760	VKGITLKYSGISIATLERLOSEFRPLPEPLTVQLLTVPGEVFPKRYKTYTFVNDVDFSMOS	819
Db	794	YKGVVLYRSGSALERIRSFSPKEPTLIQVLTV-GNALRPKIKYTYFV-----	842
QY	820	SKERATNTIQLPLHAQWYLDMEBCSSTCGAGMQRITVECRDPSGQASATCNKALKPED	879
Db	843	KKKSESFAL-PTSA-WVIEEWBCSKSCELGMORRLVECRDINGOPASEAKVKKPAS	900
QY	880	AKPCESQUCP 889	
Db	901	TRPCADHPCP 910	
RESULT 15			
AAV53899			
ID	AAV53899	standard; protein; 950 AA.	
AC	AAV53899;		
DT	13-MAR-2000	(first entry)	
DE	Amino acid sequence of a murine ADAMTS-1 protein.		
XX	GON-1; metalloproteinase; cell migration; modulator; ADAMTS-1;		
KM	metalloproteinase domain; thrombospondin domain; abnormal cell migration;		
XX	organ shaping; sterility; cancer metastasis.		
OS	Mus sp.		
PN	WO961656-A2.		
XX	02-DEC-1999.		
PF	28-MAY-1999;	99WO-US011918.	
XX	29-MAY-1998;	98US-0087170P.	
PR	13-APR-1999;	99US-0129023P.	
PA	(WISC) WISCONSIN ALUMNI RES POUND.		
PI	Kimble JE, Blueloch RH;		
XX	WPI; 2000-072633/06.		
PT	Identifying modulators of proteins containing metalloproteinase and		
XX	thrombospondin domains, potentially useful for controlling cell migration		
XX	and organ shaping.		
PS	Disclosure; Fig 1C; 60pp; English.		
XX	The present sequence represents a murine ADAMTS-1 protein. ADAMTS-1 is a		
CC	metalloproteinase. The specification describes another related		
CC	metalloproteinase, a C. elegans GON-1 protein, that lacks a transmembrane		
CC	domain and possesses a predicted metalloproteinase domain between residues		

CC 269-456. In *C. elegans* hermaphrodites, GON-1 is required for migration of  
CC two distal tip cells to produce elongated tubes, whereas in males, GON-1  
CC is required for migration of a single linker cell to produce a single  
CC elongated tube. The protein is used in the method of the invention. The  
CC specification describes a method for identifying a modulator of a protein  
CC that contains a metalloproteinase domain and a thrombospondin domain. The  
CC method comprises treating a target organism, having a developing gonadal  
CC cell that is responsive to the protein, with a test compound, and  
CC determining any change in migration or shape of the cell attributable to  
CC the test compound. The compounds identified are potential therapeutic  
CC modulators of abnormal cell migration and organ shaping, e.g. for  
CC rendering animals (specifically nematodes) sterile and for inhibiting  
CC cancer metastases  
CC  
XX  
SQ Sequence 950 AA;

Query Match 47.8%; Score 2317.5; DB 3; Length 950;  
Best Local Similarity 49.2%; Pred. No. 1,7e-155;  
Matches 450; Conservative 143; Mismatches 242; Indels 79; Gaps 21;

QY 14 LLLLLLLPLIARGAPAPPAAGQASLVVPT--RLPG-SAGELALHLSAFKGVLRJA 70  
DB 21 LLLASITMLLCARAGHGRFTE--EDEHLVLPSELRAPGHSTTRRLDAFGQHLKIQ 78  
QY 71 PDDSEFLAPEFKIELGSGSRATGER-----GLRGCFPGTVNGEPESLAASLCRGL 123  
DB 79 PDGCFLLAPGFTLQTV--GRSPGEAQHLDPTGDLACFTSGTVNGDPGSAALSLCEGV 135  
QY 124 SGSEFLDGESEFTIQOGAGSLAQPHRLQKGPAGARPLRGPWEVEYEGE---ORQE 179  
DB 136 RGARYLQGESEFTIQPAP-----GVATERLAPVPEEBSAPRQPHILRRR 181  
QY 180 RGD-----HQDSEESQEEAAGSEPPPLGA-----TSRTKRFVSEA 219  
DB 182 RGSAGACGVMDDETLPTSDRPSQNTKQWVPVDRPTPODAGKPSGPGSIRKKRFVSSP 241  
QY 220 RFVETLIVADASMAFYGADLQNHILTLMSVAARIYKPSIKNSINLMVVKLIVDEKV 279  
DB 242 RYVETMLVADQSMADFHSGSLKHLTLFSAARFYKHPISIRNSISLVVVKILVIEBK 301  
QY 280 GPEVSDNGGLTLRNFCWQRRFNQPSDRHPHRYTALILLTRONFCQGEGLCDTLGVADIG 339  
DB 302 GPEVTSNALTILRNFCWQKQNSPSDRDPHYDTALITRQDLCSH--TCDTLGMADV 360  
QY 340 TICDPNKSQSVIEDEGLQAHTTLAELGHVLSMHPDPSKPCSTRLFGPMGKHVWAPLVH 399  
DB 361 TVCDPSRSCSVIETDDG-QAFTTAHELGHVFMHPDADAKHCAASLGVSGDSHLMAWLS 419  
QY 400 LNOGLPMSPCSAMVITELLDGHDCLDAFGAALLPLTGLPGMALIYOLDQOCROIFGP 459  
DB 420 LDHSQPMSPGSAVWYTSFLDNGHCECLMDKQNPICKLPSDLPG--TLYDANRQCCFTFE 477  
QY 460 DFRHGNISADVCQAQLWC-HTDGAELPCHTKNGSLPWAADGTPCAGHLCSGSLPEE 518  
DB 478 ESKHCPD--AASCTTLMCTGTSGGLVCTKH--FPWADTSGEGKWCWCKVAKWKT 533  
QY 519 VERPKPVVDGAWPWPWEGSCRTGGGVQSHRECKDPEPONGGRYCLGRRAKYQCHT 578  
DB 534 MHPFATPVHSGWPGWGDGSRCTGGGVQYTMRECDNPVKNAGKYCEGRVRYRSNI 593  
QY 579 EECPP-DGKSPREOCCEKKNANYTDMGN--LLQWPKYAGVSPDRCKLFCARGRSE 635  
DB 594 EDCPPNNGTFRFEEQCEAHNESKASF-GNEPTVEMTPKIAGVSPDRCKLTCRAKIGY 652  
QY 636 FKVFPAKYLDGTLGPEETLALCVRGQCVKAGGDHVVDSPRKLDKGVCGKAGNSCRKXVG 695  
DB 653 FVLQPKVVDGTPGSPDSTSVCGQCVKAGCDRIIDSKKFKDKGVCGGNGSTCKXMSG 712  
QY 696 SLTPNNGYNDIVITIPAGATNIDVQKSHPGVQNDGNALAKTADGOYLINGNLASALE 755  
DB 713 IYSTRPGYHDIVITIPAGATNIEVHRNQRSGRNGSFLALAAADGTYILNGNFTLSTLE 772  
QY 756 QDILVGTLLKYSIATLERLQSFRLPEPLTVQLTLVPGHVPPKVKYTFVFNDDVF 815

DB 773 QDLYKGVTVLSSSSAALERIRSFPLKEPLTIQVLMV-GHALRPKIKTYFM----- 825  
QY 816 SMOSSKERATTNIIQPLIHQWVLGDWSECSSTGAGWQRTVECRDPGQASATCKAL 875  
DB 826 ----KKTVESFNALPTF--SEWVIEWGECSKTGSGWQRRVVQCRDINGHPASECAKEV 879  
QY 876 KPEDAKPEBSQLCP 889  
DB 880 KPASTRPCADLPCP 893

Search completed: May 7, 2004, 11:49:29  
Job time : 70 secs



Fri May 7 12:18:01 2004

us-09-989-687-4\_1.rat

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:48:10 ; Search time 23 Seconds  
(Without alignments)  
1997.701 Million cell updates/sec

Title: US-09-989-687-4

Perfect score: 4853

Sequence: 1 MFPAAPAPRLPFLILLLL.....CNKAKPEDAKPCESQLCPL 890

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTCUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3919.5	80.8	905	4	US-09-369-364A-9
2	2331.5	48.0	949	4	US-09-568-559-2
3	2328.5	48.0	967	4	US-09-130-491-2
4	2274.5	46.9	481	4	US-09-130-491-8
5	2103.5	43.3	727	4	US-09-445-023A-1
6	2093.5	43.1	727	4	US-09-445-023A-12
7	1842.5	38.0	837	4	US-09-122-126B-2
8	1842.5	38.0	837	4	US-09-634-286A-2
9	1803.5	37.2	608	4	US-09-130-491-13
10	1767.5	36.4	930	4	US-09-369-364A-2
11	1734.5	35.7	930	4	US-09-122-126B-15
12	1734.5	35.7	930	4	US-09-634-286A-13
13	1608	33.1	1882	4	US-09-369-364A-13
14	1562	32.2	874	4	US-09-369-364A-15
15	1476	30.4	2165	4	US-09-800-729-155
16	1460	30.1	551	4	US-09-130-491-16
17	1190	24.5	997	4	US-09-369-364A-7
18	1169	24.1	245	4	US-09-369-364A-11
19	1157	23.8	518	4	US-09-369-364A-22
20	1109	22.9	908	4	US-09-963-791-2
21	1088	22.4	1211	4	US-09-491-522-5
22	1070	22.0	1224	4	US-09-930-872-4
23	1054.5	21.7	757	4	US-09-963-791-24
24	1041	21.5	1205	4	US-09-491-522-11
25	1022.5	21.1	1081	4	US-09-369-364A-17
26	957	19.7	859	4	US-09-369-364A-5
27	680	14.0	589	4	US-09-963-791-12

28	625.5	12.9	438	4	US-09-963-791-22	Sequence 22, Appl
29	549	11.3	566	4	US-09-491-522-7	Sequence 7, Appl
30	495.5	10.2	1745	4	US-09-800-729-89	Sequence 89, Appl
31	492	10.1	507	4	US-09-963-791-10	Sequence 10, Appl
32	488.5	10.1	525	4	US-09-369-364A-21	Sequence 21, Appl
33	437.5	9.0	356	4	US-09-963-791-20	Sequence 20, Appl
34	425.5	8.8	468	4	US-09-963-791-6	Sequence 6, Appl
35	411.5	8.5	491	4	US-09-930-872-2	Sequence 2, Appl
36	406.5	8.4	262	4	US-09-800-729-125	Sequence 125, App
37	371	7.6	317	4	US-09-963-791-16	Sequence 16, Appl
38	352	7.3	746	4	US-09-548-797B-4	Sequence 4, Appl
39	336.5	6.9	812	4	US-09-634-098-4	Sequence 4, Appl
40	333	6.9	849	4	US-09-548-797B-6	Sequence 6, Appl
41	329.5	6.8	802	4	US-09-634-098-2	Sequence 2, Appl
42	324	6.7	775	4	US-09-786-256C-15	Sequence 15, Appl
43	324	6.7	775	4	US-09-786-256C-32	Sequence 32, Appl
44	322	6.6	787	4	US-09-548-797B-5	Sequence 5, Appl
45	309.5	6.4	1059	4	US-09-800-729-217	Sequence 217, App

#### ALIGNMENTS

RESULT 1  
US-09-369-364A-9  
; Sequence 9, Application US/09369364A  
; Patent No. 6391610  
; GENERAL INFORMATION:  
; APPLICANT: Apte, Suneel  
; APPLICANT: Hurskainen, Tina L.  
; APPLICANT: Hurskainen, Satosh  
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
; FILE REFERENCE: 26473/4007/10-30-00  
; CURRENT APPLICATION NUMBER: US/09/369,364A  
; NUMBER OF SEQ ID NOS: 31  
; SORTWARE: Patentin Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 905  
; TYPE: PRT  
; ORGANISM: Mus musculus ADAMTS-8  
US-09-369-364A-9

QY	Query Match	80.8%; Score 3919.5; DB 4; Length 905;
QY	Best Local Similarity	79.2%; Pred. No. 0; Indels 15; Gaps 4;
QY	Matches 717; Conservative 74; Mismatches 99;	
QY	1 MFPAAPAPRLPFLILLLL..LLPLARGAPAPAPAGQASSELVPTRLPGSAGELIHLIS 59	
QY	1 MRDPTTGWPLILLLL..QLPPPLVCGAPAGPTGAQASSELVPTRLPGSASELAFLIS 60	
QY	60 ARKGQFVRLADDSFLAPRFKIERLGGSGRATGGERGLRGCFEGCTNGGEPESLAAYSL 119	
QY	61 ARGGQFVRLADDSFLAPRFKIERLGGSSAAAGSGPLRGCFEGCTNGGEPESLAAYSL 120	
QY	120 CRLGSGFLDDEEFTIQPGAGSILAOPHRLQWGPAGAR.....PLRPGPW 168	
QY	121 VAGMSGFLIABEFTIQPGAGSILADPHRLQWGPAGAR.....PLRPGPW 180	
QY	169 EVETGEGQROERGDHOEDSEESQE..BEAGASPPPLGATSTYKRFVSARVETLL 226	
QY	181 EYEMGQGGQERSDNEEDKQKDEILKETEBSRVPPFGSKTSKRFVSARVETLL 240	
QY	227 VADASMAAYGADLQNHILITLMSVARIYKPSIKNSINLWVVKVLIYEDKGEVSDN 286	
QY	241 VADASMAAYGADLQNHILITLMSVARIYKPSIKNSINLWVVKVLIYEDKGEVSDN 300	
QY	287 GGLTLNFCNWRFRNPSDRPHRYDAILLITRONFCQGBELCTGLYADIGTICDPK 346	
QY	301 GGLTLNFCNWRFRNPSDRPHRYDAILLITRONFCQGBELCTGLYADIGTICDPK 360	
QY	347 SCGVIDEGLQAHILTAHETGHVLSMPHDSKPCRLRGPMKHYMAFLPHLQNTLPW 406	

```

Db      361 SCSEVIEDEGLQAAYTLAHEIGHVLSMPHDSKPCVLFSPMGKYNMADFLHVNKTLFW 420
QY      407 SPCSAMVLTBLDGGHDDCLLDAPGAALPLPTGLPGRMALYOLDQOCROI FGPDPFRHCN 466
Db      421 SPCSAYTLTBLDDGHDDCLLDAPSVLPPLTGLPGRHSTLYELDQOCROI FGPDPFRHCN 480
QY      467 TSAQDVCAQIMC-HTDGAEPPLCHTKGSLPMADGTFCGGHLCSEGSCLPESEVERPKV 525
Db      481 TSVEDICVCLCAHRRSDSEPICHTKGSLMADGTFCGGHLCSEGSCLPESEVERPKV 540
QY      526 VDDGMPWPMGSGSRCTCGGVOFSHRECKDEPDPNGRCLGRARAKYOSCHTECECPDG 585
Db      541 VDDGMPWPMGSGSRCTCGGVOFSHRECKDEPDPNGRCLGRARAKYOSCHTECECPDG 600
QY      586 KSFRQOCCEKYNAYNTMDGNLLQWPKYAGVSPDRCKLFCRANGSEPFVFEAKVID 645
Db      601 KSFRQOCCEKYNAYNTMDGNLLQWPKYAGVSPDRCKLFCRANGSEPFVFEAKVID 660
QY      646 GTLCGPETLACVRGQCVKAGCDHVDPSPKLDKCGVCGKGNCKRVSGLPTNYGN 705
Db      661 GTLCGPETLACVRGQCVKAGCDHVDPSPKLDKCGVCGKGNCKRVSGLPTNYGN 720
QY      706 DIVTIPAGATNIDVKORSHPGVQNDGNLYALKTADQYLINGNLAIISAI BODILVKGTL 765
Db      721 DIVTIPAGATNIDVKORSHPGVQNDGNLYALKTADQYLINGNLAIISAI BODILVKGTL 780
QY      766 KYSGSITLIERLOSFRPLPEPLTVOLLTVPGEVFPKVKYTFEPVNDVDFSMOSKERAT 825
Db      781 KYSGSITLIERLOSFRPLPEPLTVOLLTVPGEVFPKVKYTFEPVNDVDFSMOSKERAT 840
QY      826 TNIIOPLHAQWVLDGMSSECSSTCGAGMORRYECRDPGQASATCNKALKPEDAPCES 885
Db      841 TNIIOPLHAQWVLDGMSSECSSTCGAGMORRYECRDPGQASATCNKALKPEDAPCES 900
QY      886 QLCPL 890
Db      901 QCPPL 905

RESULT 2
US-09-568-559-2
; Sequence 2, Application US/09568559
; Patent No. 6649377
; GENERAL INFORMATION:
; APPLICANT: Klonowski, Paul
; APPLICANT: Allard, John
; APPLICANT: Heller, Renu
; APPLICANT: Van Wart, Harold
; TITLE OF INVENTION: Human Aggrecanase and Nucleic Acid
; FILE REFERENCE: ROCH-002
; CURRENT APPLICATION NUMBER: US/09/568, 559
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: 60/133,343
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 949
; TYPE: PRT
; ORGANISM: human
US-09-568-559-2

Query Match      48.0%; Score 2331.5; DB 4; Length 949;
Best Local Similarity 49.8%; Pred. No. 6e-189;
Matches 453; Conservative 145; Mismatches 249; Indels 63; Gaps 22;

```

```

QY      11 LPELILLLLLPLARGAAPRAGGQASELVVP--TLPISAGELALHLSAFSGFVLR 68
Db      15 VPTLLALLAALLAAS-DLGRPSF--EBELVVELEBRAPH-CTTLRLHAFFQQLDLE 70
QY      69 LAPDSSFLAPEFKIERLOGSGRATGG-----ERGIKGFSGTGVNGEPESLAANSLRGL 123

```

```

Db      71 LRPDSSFLAPGFTILQNV---GRKSGSETPLPENTDLACHFCYSGTVNGDPSSAALSLCEGV 127
QY      124 SGEFLDGEETITICPOGAGSLAOPHRLQRMWGPAGAPLP-----RGEMEV---ET 172
Db      128 RGAFVYLGEAFVFIQPLPAAS-----ERLATAAPEKBPALQFLLLRNKGQDVGTGCV 182
QY      173 GEGORORGDHGESESESEGESEBEAEGASEPP-----PLGATS--RTKRFVSEARFV 222
Db      183 VDDEPRFTYKAETDEDEGEDEGEDEGAQMSBODPALOGVQPTGTSIRKRRFVSSHRV 242
QY      223 ETLVADASVAAFYADLQNHILTLMSVAARIYKPSIKNSINLMVVKVILVEDEKMGPE 282
Db      243 ETLVADASVAAFYADLQNHILTLMSVAARIYKPSIKNSINLMVVKVILVEDEKMGPE 302
QY      283 VSDNGGLTIRNFCWQORRFPNPSDRHEHDTALLRKNFCGEGCDTLGVADIGTIC 342
Db      303 VTSNPAALTRNFCWQORRFPNPSDRHEHDTALLRKNFCGEGCDTLGVADIGTIC 361
QY      343 DPNKSCSVIEDEGLQAAYTLAHEIGHVLSMPHDSKPCVLFSPMGKYNMADFLHVNKTLFW 402
Db      362 DPNKSCSVIEDEGLQAAYTLAHEIGHVLSMPHDSKPCVLFSPMGKYNMADFLHVNKTLFW 421
QY      403 TLPMSPCSAMVLTBLDGGHDDCLLDAPGAALPLPTGLPGRMALYOLDQOCROI FGPDPFR 462
Db      422 TLPMSPCSAMVLTBLDGGHDDCLLDAPGAALPLPTGLPGRMALYOLDQOCROI FGPDPFR 479
QY      463 HCPNTSAQDVCAQIMC-HTDGAEPPLCHTKGSLPMADGTFCGGHLCSEGSCLPESEVER 521
Db      480 HCPNTSAQDVCAQIMC-HTDGAEPPLCHTKGSLPMADGTFCGGHLCSEGSCLPESEVER 535
QY      522 KPEVVDGMAWPMGSGSRCTCGGVOFSHRECKDEPDPNGRCLGRARAKYOSCHTECECPDG 581
Db      536 KPEVVDGMAWPMGSGSRCTCGGVOFSHRECKDEPDPNGRCLGRARAKYOSCHTECECPDG 595
QY      582 PPDGKSRQOCCEKYNAYNTMD-DGNLLQWPKYAGVSPDRCKLFCRANGSEPFVFEAKVID 639
Db      596 PPDGKSRQOCCEKYNAYNTMD-DGNLLQWPKYAGVSPDRCKLFCRANGSEPFVFEAKVID 655
QY      640 EAKVIDGLTLPETLACVRGQCVKAGCDHVDPSPKLDKCGVCGKGNCKRVSGLPTNYGN 699
Db      656 EAKVIDGLTLPETLACVRGQCVKAGCDHVDPSPKLDKCGVCGKGNCKRVSGLPTNYGN 715
QY      700 TNYGNADIVTIPAGATNIDVKORSHPGVQNDGNLYALKTADQYLINGNLAIISAI BODIL 759
Db      716 AKPGYHDITLTPGTATNLEVQKORNGSRNNGSFLAIAADGTIILINDYTLSTLEQDIM 775
QY      760 VKGTLIKTSGSIATLERLOSFRPLPEPLTVOLLTVPGEVFPKVKYTFEPVNDVDFSMOS 819
Db      776 YKGVILRYSGSAALERIRSFSPLEKEPLTIVLV-GNALRPKIKYTFV----- 824
QY      820 SKERATNIIOPLHAQWVLDGMSSECSSTCGAGMORRYECRDPGQASATCNKALKPED 879
Db      825 KKKKESFNAI-PTFSA-VWIEWEGCSKSCELGMORLVECRDINGQPAASECAKEVRKAS 882
QY      880 AKPCESQLCP 889
Db      883 TRPCADHPCP 892

RESULT 3
US-09-130-491-2
; Sequence 2, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Andrew D.J.
; APPLICANT: Holzman, Douglas A.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; PRIOR FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961

```

Fri May 7 12:18:01 2004

us-09-989-687-4\_1.ra1

Page 3

EARLIER FILING DATE: 1997-08-06  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 2  
LENGTH: 967  
TYPE: PRF  
ORGANISM: Homo sapiens  
US-09-130-491-2

Query Match 48.0%; Score 2328.5; DB 4; Length 967;  
Best Local Similarity 49.7%; Pred. No. 1.1e-188;  
Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;

11 LFFLLLLLLLLPLRGAAPAPAGGQSELVVP--TTLPGAGELALHLAFGKGYLR 68  
33 VETLLMAAALAAVS-DALGRPSR--EDELVVPLELRAFGH-GTTRLRIRAFDOQDL 88  
69 LAPDSEFLAPFKIERLGGSGRATG---ERGLRGCFSGTVNGEPELSAALSICGL 123  
89 LAPDSEFLAPFKIERLGGSGRATG---ERGLRGCFSGTVNGEPELSAALSICGL 145  
124 SGGFLDGEFTTQPOGAGGSLAQPRRLQRMGAPARPP-----RGEPEV---ET 172  
146 RGAFFYLLGAYFLQPLPAAS---ERLTAAPGEXPPAPLQFILLRNRQGVGTGCV 200  
173 GEGQROERGDHDESEESQEEAEAGSEPP-----PPLGATS-RYKRFVSEARFV 222  
201 VDDEPPTKATREDEDETEGDEGQPSQDPALQGVGCTGTGSRKRFVSHRYV 260  
223 ETLVADAMAAYGADLONHITLMSVARIYKPSININMVKYLIEDSKMPE 282  
261 ETLVADAMAAYGADLONHITLMSVARIYKPSININMVKYLIEDSKMPE 320  
283 VSDNGELTRNFCNMGORRPNOPSDRPHEDTAILLTRQFCGQGLCDTLGADIGTIC 342  
321 VTENALTLRNFNMGORRPNOPSDRPHEDTAILLTRQFCGQGLCDTLGADIGTIC 379  
343 DPNKSCSVIEDGLQAHTLAHELGHVLSMPHDSKPCRTLFGPMKHHWAPLVHLNQ 402  
380 DPNKSCSVIEDGLQAHTLAHELGHVLSMPHDSKPCRTLFGPMKHHWAPLVHLNQ 439  
403 TLPMSGSAVYLTLLDGHGDCDLPAGALPLPTGPGMALYQLDQCRQIGPDR 462  
440 SQMSPASAVMTSFDNGHGECLMDXQNPLOLBEDLPG--TSVDANQCCFTGEBDK 497  
498 HCPD--AASITCSITLWCTGSGVLCQTKH--FPWADGTSCEGSKVCINXKTRKH 553  
463 HCPNTSAODVCAQLWC-HTDGAEPICHTKNGSLPMADGTPCGPGLHCEGSLPEEVER 521  
522 PKRVVDGGAAPNPMWECSTTCGGGVQFQSHRECKDEPONGRKYCLGRAPKQSCHEHC 581  
554 FDTFFGSGWGMWPMWDCSTTCGGGVQFQSHRECKDEPONGRKYCLGRAPKQSCHEHC 613  
582 PP-DGKSFREOQCEKNAAYNYTDM-DGNLLQWPKTAGVSPRDRCKLFCRARGSEFKV 639  
614 PDNNGKTFREOQCEKNAAYNYTDM-DGNLLQWPKTAGVSPRDRCKLFCRARGSEFKV 673  
640 EAKYIDGTLGPEETLAI CVRGQCVKAGCDHVVDSPKLDKGVCGGKNSCRKVSGLTP 699  
674 QPKVYDGTSPDSTVCGQCVKAGCDHVVDSPKLDKGVCGGKNSCRKVSGLTP 733  
700 TNGYNDIVITIPAGATIDVYKORSHPGVQNDGNVYLAKTADGOYLNGNLALSAIEDIL 759  
734 AKGKHITITIPGATIDVYKORSHPGVQNDGNVYLAKTADGOYLNGNLALSAIEDIL 793  
760 VKGTLIKSGSIATLERLQSFRLPEPLTVQLLTPGEVFPKVKYTFPVNDVDFSMOS 819  
794 YKGVVLAAYSSGSAALERIRSFSLPEPLTVQLLTPGEVFPKVKYTFPVNDVDFSMOS 842  
820 SKERATNIIQPLHAQWVLDGMEGSCSTGAGMORRYVCRDPGQASATCNKALPED 879  
843 KKKKESFNAL-PTFSA-VVLEMGESKSCLEGMORRLVBCRDINQPSACEKAYKPS 900

QY 880 AKPCESOLCP 889  
DB 901 TRPCADHPCP 910

RESULT 4  
US-09-130-491-8  
Sequence 8, Application US/09130491  
Patent No. 6416974  
GENERAL INFORMATION:  
APPLICANT: Holtzman, Douglas A.  
APPLICANT: Goodheart, Andrew D.J.  
TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83  
FILE REFERENCE: 09404/041001  
CURRENT APPLICATION NUMBER: US/09/130,491  
EARLIER FILING DATE: 1998-08-07  
EARLIER APPLICATION NUMBER: US 60/054,108  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: US 60/054,961  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 8  
LENGTH: 481  
TYPE: PRF  
ORGANISM: Rattus rattus  
US-09-130-491-8

Query Match 46.9%; Score 2274.5; DB 4; Length 481;  
Best Local Similarity 82.1%; Pred. No. 1.5e-184;  
Matches 403; Conservative 35; Mismatches 40; Indels 13; Gaps 2;

401 NOTLPMSPGSAVYLTLLDGHGDCDLPAGALPLPTGPGMALYQLDQCRQIGPDR 460  
3 NKTLPMSGSAVYLTLLDGHGDCDLPAGALPLPTGPGMALYQLDQCRQIGPDR 50  
461 FRHCNPTSAODVCAQLWC-HTDGAEPICHTKNGSLPMADGTPCGPGLHCEGSLPEEVER 519  
51 FRHCNPTSAODVCAQLWC-HTDGAEPICHTKNGSLPMADGTPCGPGLHCEGSLPEEVER 550  
520 ERPPRVVDGGAAPNPMWECSTTCGGGVQFQSHRECKDEPONGRKYCLGRAPKQSCHEHC 579  
111 ENPRAVVDGGAAPNPMWECSTTCGGGVQFQSHRECKDEPONGRKYCLGRAPKQSCHEHC 610  
580 ECPBDKSFREOQCEKNAAYNYTDM-DGNLLQWPKTAGVSPRDRCKLFCRARGSEFKV 639  
171 ECPBDKSFREOQCEKNAAYNYTDM-DGNLLQWPKTAGVSPRDRCKLFCRARGSEFKV 670  
640 EAKYIDGTLGPEETLAI CVRGQCVKAGCDHVVDSPKLDKGVCGGKNSCRKVSGLTP 699  
231 ETKYIDGTLGPEETLAI CVRGQCVKAGCDHVVDSPKLDKGVCGGKNSCRKVSGLTP 730  
700 TNGYNDIVITIPAGATIDVYKORSHPGVQNDGNVYLAKTADGOYLNGNLALSAIEDIL 759  
291 FSYGNDIVITIPAGATIDVYKORSHPGVQNDGNVYLAKTADGOYLNGNLALSAIEDIL 800  
760 VKGTLIKSGSIATLERLQSFRLPEPLTVQLLTPGEVFPKVKYTFPVNDVDFSMOS 819  
351 MKGTLIKSGSIATLERLQSFRLPEPLTVQLLTPGEVFPKVKYTFPVNDVDFSMOS 840  
820 SKERATNIIQPLHAQWVLDGMEGSCSTGAGMORRYVCRDPGQASATCNKALPED 879  
411 SKERATNIIQPLHAQWVLDGMEGSCSTGAGMORRYVCRDPGQASATCNKALPED 900  
880 AKPCESOLCP 890  
471 AKPCESOLCP 481

RESULT 5  
US-09-445-023A-1  
Sequence 1, Application US/09445023A  
Patent No. 656588

```

GENERAL INFORMATION:
APPLICANT: Hirose, Kunitaka
APPLICANT: Inoguchi, Eiji
APPLICANT: Hakozaaki, Michinori
APPLICANT: Ishida, Keiko
APPLICANT: Ishida, Yukako
APPLICANT: Matsushima, Kouji
APPLICANT: Kuno, Kouji
TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
FILE REFERENCE: 057092
CURRENT FILING DATE: 1999-12-03
PRIORITY FILING DATE: 1997-06-03
PRIORITY FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin version 3.0
SEQ ID NO 1
LENGTH: 727
TYPE: PRT
ORGANISM: Homo sapiens
US-09-445-023A-1

Query Match      43.3%; Score 2103.5; DB 4; Length 727;
Best Local Similarity 54.8%; Pred. No. 9.2e-170;
Matches 374; Conservative 121; Mismatches 164; Indels 23; Gaps 11;

211 RTKRPFSEARFVEFLVADASMAAFYGADIQNHITLMSVAARIYGHPSIKNSIMLVK 270
Db |K|K|K|F|V|S|P|R|V|E|T|M|V|A|D|S|M|A|F|H|G|S|G|K|H|Y|L|T|F|S|V|A|A|R|Y|G|H|P|S|I|N|S|I|N|V|V|K| 68
9 271 V|I|V|E|D|E|K|W|G|E|V|S|D|N|G|L|T|I|R|N|F|C|W|O|R|R|E|N|O|P|S|R|H|E|H|Y|T|A|I|L|T|R|O|N|F|C|G|E|G|C 330
Qy |I|V|H|D|E|Q|K|P|E|V|T|S|N|A|L|T|I|R|N|F|C|W|O|K|O|N|P|S|R|D|A|H|Y|T|A|I|L|T|R|O|D|C|G|Q|-T|C 127
Db 69 |I|V|H|D|E|Q|K|P|E|V|T|S|N|A|L|T|I|R|N|F|C|W|O|K|O|N|P|S|R|D|A|H|Y|T|A|I|L|T|R|O|D|C|G|Q|-T|C 127
Qy 331 D|T|G|V|A|D|I|G|T|C|D|P|N|K|S|C|V|E|D|G|L|Q|A|A|T|L|A|H|E|G|H|V|L|S|M|P|H|D|S|K|E|P|T|R|L|F|G|M|G|K 390
Db 128 D|T|G|V|A|D|I|G|T|C|D|P|N|K|S|C|V|E|D|G|L|Q|A|A|T|L|A|H|E|G|H|V|L|S|M|P|H|D|S|K|E|P|T|R|L|F|G|M|G|K 390
Qy 391 H|V|A|P|L|V|H|N|O|L|P|W|S|P|C|S|A|N|Y|L|T|E|L|D|G|H|D|C|L|D|A|P|A|L|P|L|P|T|G|I|P|G|M|A|L|Y|O|D 450
Db 188 H|V|A|P|L|V|H|N|O|L|P|W|S|P|C|S|A|N|Y|L|T|E|L|D|G|H|D|C|L|D|A|P|A|L|P|L|P|T|G|I|P|G|M|A|L|Y|O|D 450
Qy 451 Q|Q|C|Q|I|G|P|P|R|H|C|P|N|T|S|A|D|V|C|A|Q|L|W|C|-H|T|D|A|E|P|L|C|H|T|N|G|S|L|P|W|A|D|T|P|G|P|H|L|S 509
Db 246 R|Q|C|F|T|G|E|D|S|K|H|C|P|D|-A|A|S|T|C|T|I|W|C|T|G|S|G|L|V|C|Q|T|H|-F|P|W|A|D|T|S|C|G|E|G|K|C|I 301
Qy 510 E|G|S|C|L|P|E|E|V|R|P|K|V|V|D|G|W|A|P|W|G|E|C|S|R|T|C|G|G|V|Q|F|S|H|E|C|K|D|P|E|P|O|N|G|R|Y|C|L|A|R 569
Db 302 N|G|K|V|N|K|T|R|K|H|P|T|P|H|G|S|G|W|P|G|P|W|D|C|S|R|T|C|G|G|V|Q|Y|T|R|E|C|O|N|P|R|V|K|N|G|K|Y|C|E|G|K 361
Qy 570 R|A|K|Y|Q|S|H|E|C|P|-D|E|K|S|F|E|O|C|E|K|N|A|N|Y|T|D|M|-D|G|L|O|W|P|R|Y|K|A|V|S|R|D|R|K|L|F 627
Db 362 R|V|R|R|S|C|L|E|D|C|P|N|N|G|T|F|E|E|O|C|E|A|N|E|R|S|K|A|S|F|G|S|G|A|V|E|I|P|R|A|V|S|P|R|D|R|K|L| 421
Qy 628 C|R|A|G|R|S|E|F|K|Y|E|A|K|V|D|I|G|T|C|G|E|T|A|I|C|V|R|G|C|V|A|G|C|H|V|V|D|S|R|K|L|D|K|G|V|C|G|K 687
Db 422 C|Q|A|G|I|G|F|V|L|Q|R|K|V|D|G|T|C|P|S|T|S|V|C|V|Q|C|V|A|G|C|R|I|D|S|K|K|R|D|K|G|V|C|G|G|G 481
Qy 688 N|S|C|R|V|S|G|L|P|T|V|Y|G|N|D|I|V|T|P|A|G|A|T|N|D|V|K|O|R|S|H|P|G|V|O|N|D|G|N|Y|L|A|L|K|T|A|D|G|Q|Y|I|L|G 747
Db 482 S|T|C|K|I|S|G|S|V|S|A|P|G|Y|H|D|I|V|T|P|T|G|A|T|N|E|V|K|O|R|N|O|R|G|R|N|N|S|F|L|A|I|R|A|A|D|G|Y|I|L|G 541
Qy 748 N|L|A|S|A|I|E|D|I|V|G|T|I|K|Y|S|G|S|A|T|E|R|L|O|S|F|R|L|P|E|P|L|T|V|Q|L|L|T|V|P|G|E|V|F|P|K|V|Y|T|F 807
Db 542 D|Y|T|S|T|E|Q|D|I|N|K|V|L|R|Y|G|S|S|A|L|E|R|I|S|F|S|P|L|E|P|L|I|Q|V|L|T|Y|-G|N|A|R|K|I|K|Y|T| 600
Qy 808 F|V|P|N|D|V|D|S|Q|S|K|R|A|T|N|I|Q|P|L|H|A|Q|V|L|G|W|S|B|C|S|T|C|G|A|G|P|R|Y|E|C|H|D|P|S|G|Q|A 867
Db 601 F|V|-----K|K|K|E|S|F|N|A|I|-P|T|F|S|A|-V|I|E|W|G|E|C|S|C|E|I|G|W|R|R|V|E|C|R|D|I|N|G|P 648
Qy 868 S|A|T|C|K|A|K|P|E|D|A|P|C|S|Q|L|C|P 889

```

```

Db 649 A|S|E|C|A|K|E|V|K|P|A|S|T|R|P|C|A|D|H|P|C 670

RESULT 6
US-09-445-023A-12
Sequence 12, Application US/09445023A
Patent No. 6565858
GENERAL INFORMATION:
APPLICANT: Hirose, Kunitaka
APPLICANT: Inoguchi, Eiji
APPLICANT: Hakozaaki, Michinori
APPLICANT: Ishida, Keiko
APPLICANT: Ishida, Yukako
APPLICANT: Matsushima, Kouji
APPLICANT: Kuno, Kouji
TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
FILE REFERENCE: 057092
CURRENT FILING DATE: 1999-12-03
PRIORITY FILING DATE: 1997-06-03
PRIORITY FILING DATE: 1997-06-03
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin version 3.0
SEQ ID NO 12
LENGTH: 727
TYPE: PRT
ORGANISM: Mus sp.
US-09-445-023A-12

Query Match      43.1%; Score 2093.5; DB 4; Length 727;
Best Local Similarity 54.0%; Pred. No. 6.5e-169;
Matches 374; Conservative 124; Mismatches 168; Indels 27; Gaps 12;

201 E|P|P|L|G|T|S|R|K|P|V|S|E|A|R|V|E|T|L|V|A|D|S|M|A|F|Y|G|A|D|I|Q|N|H|I|T|L|S|V|A|A|R|Y|G|H|P|S|I 260
Db 1  E|P|S|G|P|-G|S|I|R|K|F|V|S|P|R|V|E|T|M|V|A|D|S|M|A|F|H|G|S|G|K|H|Y|L|T|F|S|V|A|A|R|Y|G|H|P|S|I 58
Qy 261 K|N|S|I|N|V|V|K|L|I|Y|E|D|E|K|W|G|E|V|S|D|N|G|L|T|I|R|N|F|C|W|O|R|R|E|N|O|P|S|R|H|E|H|Y|T|A|I|L|T|R 320
Db 59  R|N|S|I|S|V|V|K|L|I|Y|E|D|E|K|W|G|E|V|S|N|A|L|T|I|R|N|F|C|W|O|K|O|N|P|S|R|D|A|H|Y|T|A|I|L|T|R 118
Qy 321 Q|N|F|C|Q|E|G|L|C|D|T|G|V|A|D|I|G|T|C|D|P|N|K|S|C|V|E|D|G|L|Q|A|A|T|L|A|H|E|G|H|V|L|S|M|P|H|D|S|K|E|P|C 380
Db 119 Q|D|L|G|S|H|-T|C|D|T|G|V|A|D|I|G|T|C|D|P|N|K|S|C|V|E|D|G|L|Q|A|A|T|L|A|H|E|G|H|V|L|S|M|P|H|D|S|K|E|P|C 177
Qy 381 T|R|L|P|G|M|K|H|Y|N|A|R|P|V|H|N|O|L|P|W|S|P|C|S|A|N|Y|L|T|E|L|D|G|H|D|C|L|D|A|P|A|L|P|L|P|T|G|I|P|G|L 440
Db 178 A|S|I|N|V|S|D|S|H|L|S|M|S|L|D|H|S|Q|W|S|P|C|S|A|N|Y|L|T|S|F|L|D|N|G|H|G|E|C|L|D|K|P|Q|N|P|I|K|L|P|E|D|L 237
Qy 441 P|G|R|N|A|L|Y|O|L|D|Q|C|R|O|I|F|P|D|P|R|H|C|P|N|T|S|A|D|V|C|A|Q|L|W|C|-H|T|D|A|E|P|L|C|H|T|N|G|S|L|P|W|A|D|G 499
Db 238 P|G|-T|L|Y|A|N|A|Q|C|P|T|P|E|E|S|K|H|C|P|D|-A|A|S|T|C|T|I|W|C|T|G|S|G|L|V|C|Q|T|H|-F|P|W|A|D|G 291
Qy 500 T|P|C|G|P|H|L|C|S|B|G|S|C|L|P|E|E|V|R|P|K|V|V|D|G|W|A|P|W|G|E|C|S|R|T|C|G|G|V|Q|F|S|H|E|C|K|D|P|E|P 559
Db 292 T|S|C|G|E|G|K|V|C|S|G|K|C|N|K|I|D|M|H|F|A|P|V|G|S|W|P|G|P|W|D|C|R|T|C|G|G|V|Q|Y|T|R|E|C|D|N|P| 351
Qy 559 O|N|G|R|Y|C|G|R|A|K|Y|O|S|H|T|E|C|P|-D|E|K|S|F|E|O|C|E|K|N|A|N|Y|T|D|M|D|N|-L|L|O|W|P|R|Y|K|A|V|S|R|D|R|K|L|F 616
Db 352 K|N|G|K|Y|C|E|G|K|R|Y|R|Y|S|C|N|I|E|D|C|P|N|N|G|T|F|E|E|O|C|E|A|N|E|R|S|K|A|S|-G|N|E|P|V|E|T|P|K|Y|A 410
Qy 617 G|V|S|P|R|D|R|C|L|P|R|A|R|S|E|F|K|Y|E|A|K|V|D|I|G|T|C|G|E|T|A|I|C|V|R|G|C|V|A|G|C|H|V|V|D|S|R|K 676
Db 411 G|V|S|P|R|D|R|C|L|P|R|A|R|S|E|F|K|Y|E|A|K|V|D|I|G|T|C|G|E|T|A|I|C|V|R|G|C|V|A|G|C|H|V|V|D|S|R|K 470
Qy 677 L|R|K|G|V|C|G|K|G|N|S|C|R|K|V|S|G|L|P|T|V|Y|G|N|D|I|V|T|P|A|G|A|T|N|D|V|K|O|R|S|H|P|G|V|O|N|D|G|N|Y|L|A|L 736
Db 471 F|D|K|G|V|C|G|K|G|N|S|C|R|K|G|N|G|Y|T|S|T|R|P|G|Y|H|D|I|V|T|P|A|G|A|T|N|E|V|K|R|N|O|R|G|R|N|N|S|F|L|A|I 530
Qy 737 K|T|A|D|Q|Y|I|L|N|G|L|A|S|A|I|E|D|I|V|G|T|I|K|Y|S|G|S|A|T|E|R|L|O|S|F|R|L|P|E|P|L|T|V|Q|L|L|T|P|G 796
Db 531 R|A|D|G|Y|I|L|N|G|F|L|T|S|T|E|O|D|L|Y|K|G|I|V|L|R|Y|G|S|S|A|L|E|R|I|S|F|S|P|L|E|P|L|I|Q|V|L|N|Y|-G 589

```

Fri May 7 12:18:01 2004

us-09-989-687-4\_1.rail

Page 5

Cy	79	EVPPEKVKYKFEEFVNVDVDSMOSKRATNTIIOPLLAQWLGLPMSSCSSTCGAQORR	856
		::: :	
Dδ	590	HALRPKKTFYFM-----KKRTSEFNAPIPTF--SEWIIEMWGECSTKTCSGWQR	637
Cy	857	TVECRDPSGQSASATCNKALKPEDAKPCESQLCP	889
		:: :: :	
Dδ	638	VVGCRDLNGHNPASECAKEVFAPASTRCADLPCEP	670

```

RESULT 7
US-09-122-126B-2
; Sequence 2, Application US/09122126B
; Patent No. 6451575
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909
; CURRENT APPLICATION NUMBER: US/09/122,126B
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 837
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-122-126B-2

```

Query Match	38.0%	Score 1842.5	DB 4	Length 837
Best Local Similarity	45.0%	Prod. No. 1.6e-147		
Matches 375		Mismatches 130	Indels 85	Gaps 24

```

QY 5 PPAAP-RMPLPILLILLIL-LLPLARAPAPRAPAGQASLWVTRPLPGSA-----GELA 55
Db 27 PIVPLPSMWLWLLILLILASLLPSARLASPIF---PEBELVPEPKLNGSVLPSSGAPARLL 82
QY 56 LHLISAFKGFVRLIAPDDSEFLAPEEFIERLIGSGGATGGERGLRGCFFSGVNGESPESIA 115
Db 83 CRLQAFGETLILLEBEDSGVQVEGLTVQLGQAPELLGSAE--PGYILITGINDPESIVA 140
QY 116 AVSLCRG-ISSGFLLDGEEFTIQ-----PGAGSGLAOPHLQKWRPAPPLRGPBW 168
Db 141 SLHMDGALLGLVQYKGAELHLQPLEGCTPNSGGPGA--HILRRKSPASG---QGPWC 194
QY 165 EYETGEGCQOEKDHODEBESSEQBEEAAGSAPPLATSS---RTGRFVSEARFVET 224
Db 195 NPKYA-----PLGSPSPRRAPRASFASLSRFVT 222
QY 225 LTVADASMAFYGADLQNLILTLMSVAAIYKHPISIKNSINIMVYKVLIVDEKMGPEVS 284
Db 223 LVYADDKMAFPGAGLKRKYLITVMAAAXAKFAPSIRNVSLSVTRTLVILIGSEEGEPQVG 282
QY 285 DNGGLTLRNFCMQRFRNPQSDHPEPHYDALILTEONPCGSGCLDPLGVADIGTICP 344
Db 283 PSAQQLTRSCAWQRGINTPEBDDPHFIDALLFTQDCLG-VSTCDILGMDVGVICP 341
QY 345 NKSCSYIYEDGLOAAHTLIAHELGHVLSMPHDSKPCSTRLEGPWG-KHHYAPLPLEVHNOT 403
Db 342 ARCAIYEDDGLQSAFTAAHELGHVFMHLDNKKPCISLNGPFLSTRHWAPMAHAYDEB 401
QY 404 LPMSPGSAMYLTELLDGGHGDCLLDAPGAALLPLPTGIPGMALYQLDQOCROILFGDFFH 463
Db 402 EPMSPSCARFITDLDNGYGHCLLDKEAPLHLPVTFPGD--YDARDQQLTFGGPSRR 459
QY 464 CPMTSADQVACOLWC--HTDGAEPICHTKNGSLPMAADGTPCGGPHLCSGSCULPEEBEVR 521
Db 460 CPQULPPP--CALMCSGHLNG-HAMCOTKHS--PMAHGTPCGPQAQCMGRCILHMDQLQD 514
QY 522 PKPVPVDGMAPMWGWGECSRTCGGQVQFSHECKDPEPQNGRYCLGRAPKYQSCHEEC 581
Db 515 FNIPQAGMGWGMGWDGCSRTCGGQVQFSSSDCTCRRPVBRNGKXCEGRARRFRSCTEDC 574
QY 582 PP-DGKSFREBQCGCKNYANVYTDMDGNL---LQWVPKXAGVSPDRCKLFCBARGRSEFK 637

```

```

Db      575  PTGSALTFFEEBOOAAVN--HRTDLEKSPFGPMDMVPRKTYGTGAHPDQCKLTCQARALGY 632
Qy      638  VFEAKYIDSLTSGPETHALCVARGQCVKAGCDHVVDSPRKLDKCVGCGKANSCKRYISGL 697
Db      633  VLEPRVYDGTPCSPDSSSVCYGRCTIHAGCDRIIGSKKKFKCMVCGADSGSCSKQSGSF 692
Qy      698  TPTNYGNDIVTIPACAINIDVKQSHGPGVONDGNATLAKTAQDQYLLNGLNLSALED 757
Db      693  RKFRYGNWNVITIPATHILVRQCGNG--HRSYIALAKLPDSEVALNGEYTLMPSPTD 750
Qy      758  ILVKGTT-LKYSGLITLERLOSFPPLDEPLTVQLLTPGSEVFPFKRYTFEVP 810
Db      751  VWLPGAVSLRKSGATTASETLGHQPLAQPLTLQYL-VAGNPQDTRLRKYSFVP 803

```

RESULT 8  
US-09-634-286A-2  
; Sequence 2, Application US/09634286A

1 APPLICANT: Bristol-Myers Squibb Company  
2 TITLE OF INVENTION: AGREONIN DEGRADING METALLO PROTEASES  
3 FILE REFERENCE: DM6909A  
4 CURRENT APPLICATION NUMBER: US/09/534,286A  
5 CURRENT FILING DATE: 2000-08-03  
6 NUMBER OF SEQ ID NOS: 21  
7 SOFTWARE: PatentIn version 3.0

/ SEQ ID NO 2  
 ; LENGTH: 837  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ;  
 US-09-634-286A-2

Query Match	38.0%	Score 1842.5	DB 4	Length 837
Best Local Similarity	45.0%	Pred. No. 1.6e-147		
Matches 373	Conservative 130	Mismatches 244	Indels 85	Gaps 24

```

QY      5 PAAAP-RNLPFLILLILL-LPLPARAPARPAAGGASLWVTRIPGSA-----GBLA 55
Dh      27 PVPPLSWLWMLLLILLASLIPBARLASLP-----REEVIVPEKINGSVLPSSGAPARLL 82
QY      56 LHLISAFKGFVLRILAPDPSFLAPEFKIRLGGSGRATGSGRGLRGCFFSGIVNGPESTIA 115
Dh      83 CRLQAFGETLLELDQDSGVQVEGLTYQYLIGAPRLLGGAE--PETYLTGTINGDPESSVA 140
QY      116 AVSLCRG-LSGSFLLDGEBFTIQ-----PGAGGSLAOPHRLQWGPAGAPFLPRGPBW 168
Dh      141 SLHMGGALLGLQYRGAEHLIQPLEGGTPNSAGBPGA--HLIRKKSPPSG---QGPPWC 194
QY      169 EYETGGCQGRGHDHGEDESEESQEBAGASEPPLCATG-----RTRRFVSEARFVET 224
Dh      195 NVKA-----PLGSPSPRRRARPRASLSFVET 222
QY      225 LTVADASMAAFYGADLQNHILTLMSVAARIYHPSTIKSINIMVYKVLIVEBEKMPPEVS 284
Dh      223 LVVADDKMAAFGAGLKRYYLITVMAAAAFKHPGSRINRVSVLVVTRLVITLGGSEBTPVG 282
QY      285 DNGGLTLRNFCWQRFPNOPSDBRPHNYTALLLRONCGQEGCDTLGVADIGTICDP 344
Dh      283 PSAQTLRSCAMQORBLNTPEDSDPHFTALLFTKQDCG--VSTCDTLGMADVGVICDP 341
QY      345 NKSCSVIIEBGOAAATTLAHELGHYLSMHDISKCPTRLFGPWC--KHHVAPLVEVLIOT 403
Dh      342 ARSCALVEBDGLOSATLAHELGHYVNMJHDSKCPSTLNGPLSTRHYMAPVMAHVDEP 401
QY      404 LPMSPCSAMYLTELIDGCHGDCLLDAPGALLPLPTLPGRMALYOLDQOCROIFGSDPRH 463
Dh      402 EPMSPSCARFITDLFNGYGHCLIDRBEAPLHPVTFPKGD--YDADRQOQLTFGDSRH 459
QY      464 CPMTSADYQAQWMC--HIDGAEPLCHTNKGLPMADGTPCGGPHLCSBGSCLPEEVER 521
Dh      460 CPQLPBP--CAALWCSGHING-HAMQTKHS--PMADGTPCGSAQCMGSRCLHMQLOLD 514

```

QY 522 PKRVNDGMAPMGWECSTRTCGGVQFOSHECKDEPEONGGRYCLGRAPXYOSCHTEBC 581  
 DB 515 FNIPQAGMGWPGWGCSTRTCGGVQFOSHECKDEPEONGGRYCLGRAPXYOSCHTEBC 574  
 QY 582 PP-DGKSFREOQCEKXNANYTMDGNTL--LQWPKYAGVSPDRCKLFCBARGRSEPK 637  
 DB 575 PTGSALTFRREOQCAAYV--HRTDLFKSPFGPMDMVPRYTGVAPODQCKLTCQARALGYXY 632  
 QY 638 VFAKATIDTLCGPBETALICVRGOCVAGCDHVDSPRLDKGVGCGKSGKSRKVSGL 697  
 DB 633 VLBPRVVDGTPCSPDSSVCVQGRCHTAGCDRLTIGSKKFKDKCMVCGDGSCKSGSF 692  
 QY 698 TPTNYGNDIVTTPAGATNIDVKQSHPGVQNDGNYLALKTADGOYLNGNLALISAIBOD 757  
 DB 693 RKPRYGNVNTTIPAGATHILVQGNPG--HRSYIALTLPGSYALNGEYTIMSPFD 750  
 QY 758 ILVAGIT-LKYSGLATLERLQSFRLPEPLTVQLTVPGEVPPPKYKTFEP 810  
 DB 751 VVLPFAVSLRYSGATASETLSGHGFLAQLTLQVL-VAGNPOTDRLRYSPFVP 803

RESULT 9  
 US-09-130-491-13  
 ; Sequence 13, Application US/09130491  
 ; Patent No. 6416974  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Holzman, Douglas A.  
 ; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83  
 ; FILE REFERENCE: 09404/041001  
 ; CURRENT APPLICATION NUMBER: US/09/130,491  
 ; CURRENT FILING DATE: 1998-08-07  
 ; EARLIER APPLICATION NUMBER: US 60/058,108  
 ; EARLIER FILING DATE: 1997-09-05  
 ; EARLIER APPLICATION NUMBER: US 60/054,961  
 ; EARLIER FILING DATE: 1997-08-06  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 13  
 ; LENGTH: 608  
 ; TYPE: PR1  
 ; ORGANISM: Mus musculus  
 ; US-09-130-491-13

Query Match 37.2%; Score 1803.5; DB 4; Length 608;  
 Best Local Similarity 52.2%; Pred. No. 2,1e-144; Indels 25; Gaps 12;  
 Matches 320; Conservative 113; Mismatches 155;  
 QY 271 VLIWDEKMGPEVSDNGGLTLRNFQNRFPQSDRHPHEDTALITLQNFQCGDEGLC 330  
 DB 4 ILVHDEKMGPEVSDNGGLTLRNFQNRFPQSDRHPHEDTALITLQNFQCGDEGLC 62  
 QY 331 DTLGVADIGTICDPNKSQSVIEDBGLQAHTLAHLGHVLSMHPDSDKPCRTLFGPMKHA 390  
 DB 63 DTLGVADIGTICDPNKSQSVIEDBGLQAHTLAHLGHVLSMHPDSDKPCRTLFGPMKHA 122  
 QY 391 HVMALFLVHLNQTLPMSPCSAWYLTLDGGHDCILDAFGAALPLFTGLFGPMALYOLD 450  
 DB 123 HVMALFLVHLNQTLPMSPCSAWYLTLDGGHDCILDAFGAALPLFTGLFGPMALYOLD 180  
 QY 451 QOCROIFGDPFRHCNTSAQDVQALWC-HTDGAPLCHTXGSLPMADGTPCPGHLCS 509  
 DB 181 RQCGTTFGSDSGHCDP--AATGCTLMCTGTSGVLYQVQIKR--FPWADGTSCEBGRKCI 236  
 QY 510 BEGCLPEEVEVERPKVVDGMAPMGWECSTRTCGGVQFOSHECKDEPEONGGRYCLGR 569  
 DB 237 NKCVCNKTDRKAFDTPFRSGWGMGWPMDCKRTCCGGVQYIMREBDNVPRKMGKCYCEK 296  
 QY 570 RAKYOSCHTEBCPP-DGKSFREOQCEKXNANYTMD-DGNLQWVPKAYGSPDRCLTF 627  
 DB 297 RYKYSRCSNLEDCPDNNGKTFREOQCEAHNEFSKASFGSGPAVEMLPKAYGSPDRCLTF 356

QY 628 CPARGSEPKYFEAKYIDTLCGPBETALICVRGOCVAGCDHVDSPRLDKGVGCGK 687  
 DB 357 CQAKGIGYFFFLQPKVVDGTPCSPDSSVCVQGRCHTAGCDRLTIGSKKFKDKCMVCGDGSCKSGSF 416  
 QY 688 NSCRKVSGLTPTNYGNDIVTTPAGATNIDVKQSHPGVQNDGNYLALKTADGOYLNGNLALISAIBOD 747  
 DB 417 STCKKISGVSATPAGHDIITIPIGATNIEVQNRGSRKNSFLAIAADQTYLNG 476  
 QY 748 NLASAIEDILVKTILKTSGLATLERLQSFRLPEPLTVQLTVPGEVPPPKYKTFEP 807  
 DB 477 DYTSLTEODIMYGKVLVYSGSSAALERTSFSPLKEPLTIVTV-GNALRPKITYY 535  
 QY 808 FVPNDVDSMSSKERRATNTIQLPLLAOWVLGDMSECSSTCGAGMORATVECDPESG-- 865  
 DB 536 FV-----KKKSFNAI-PTPSA-VWIEBWSCKTCKGKYGKRLKLSHDG 583  
 QY 866 QASATCNKALKE 878  
 DB 584 LSHSCDPLKKPK 596

RESULT 10  
 US-09-369-364A-2  
 ; Sequence 2, Application US/09369364A  
 ; Patent No. 6391610  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Apte, Suneel  
 ; APPLICANT: Hirschman, Tina L.  
 ; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
 ; FILE REFERENCE: 26473/4007/10-30-00  
 ; CURRENT APPLICATION NUMBER: US/09/369,364A  
 ; CURRENT FILING DATE: 1999-08-06  
 ; NUMBER OF SEQ ID NOS: 31  
 ; SOFTWARE: Patentin Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 930  
 ; TYPE: PR1  
 ; ORGANISM: mus musculus ADAMTS-5  
 ; US-09-369-364A-2

Query Match 36.4%; Score 1767.5; DB 4; Length 930;  
 Best Local Similarity 40.2%; Pred. No. 4.5e-141;  
 Matches 380; Conservative 125; Mismatches 355; Indels 85; Gaps 19;  
 QY 10 WLPFLILLILL--LLPLARGA-----RPAAGGQASLVYP----- 44  
 DB 5 WSLILLILLILLASCLSLAADSPPAAPADQKTRCPQAAAAAEPDQGEETREGRGLQP 64  
 QY 45 -----TRPGSAGELALHLSARGKGVRLAPDPSFLAEPFKIERLGSGR 90  
 DB 65 LAGQRRSGGLVHNDQLVSGGKGVLYVYAGRRFLDLERDIT-VGAAGSLVTAGGGLS 123  
 QY 91 ATGGERGLRGCFSGTVNGEPESLAVALCRGLSSFLIDGEEFTIQPGAGGSLAQPHR 150  
 DB 124 ASSGRHG--HCFYRGTVDSPRSLAVFDLCGLDGFVAKHARYLTKPL--LGSMAEYER 180  
 QY 151 LQRMGPAGAR-----PLPRGPEVEVEGEQROREGRHOESBESQ----- 192  
 DB 181 I--YGDSSRLILHVYNRBSFSEALPPRASCTTPASPSGPDSPVHNSRRRSALAPL 238  
 QY 193 -EEBAEGASEPPPLCATGRTFRFVSEAFVETLLVADASMAFYGADIQNHILTLMSVA 251  
 DB 239 LHSASAPSGNAGPQWMMRRRRRSISRAGVELLLVADSMARYGRGLQHYLLITWASTA 298  
 QY 252 ARTYKPSIKNSINLNVYKVLVEDEKMGPEVSDNGGLTLRNFQNRFPQSDRHPH 311  
 DB 299 NRLYSHASIEHNLRLAVVYVYVLTDDTSLVSKNAATLKNPCKKQHQHNLGDDHEH 358  
 QY 312 YDFAILLTRONFCGEGCLDITLGVADIGTICDPNKSQSVIEDBGLQAHTLAHLGHVLS 371  
 DB 359 YDAIILFTBEDLGHHS--CDTLGMADVGTICSPERSCAVIEDDGLAAATVAHEIGHLLG 417

```

QY 372 MPRHDSKPCSTRLPGRMKGHVMAALFYHNLQTLPMWSCSAMLYLLELLDDGSHGDCILDAPG 431
Db 418 LSHDSDKFCCEENFTGTTEDKRLMSSILTSIDASKRWSKCTSAITTEFLPDHGNCLEDLPR 477
QY 432 AALPLPLPGLPRMALYOLDOOQCRQILFGDFHNCENTSADODCAOLMCH-IDGAEPLOCHTK 490
Db 478 KQIGPEELPQ--TYDATQOONLTFGEYVCP---GMDCALMLCAVVRQGMVCLTK 532
QY 491 NGLPMADGTPCGRGLCSSEGSCLPEEVEVERPKVVDGMAVPMQMECESTCGGYQFS 550
Db 533 --KLPAVEGTDCGKGRVCLQCKCYDKTKKRYTSTSHGMGSMGPMQCSRS CGGYQFA 590
QY 551 HRECKDEPPONGRYCLGRAPKYOQSCHEECPPOGKSFRBOOCEKRYANVYTDMDG--NL 608
Db 591 YRHCHNPAFNRSGRYCTGKRAIYRSCSVTPCEPENGKSRHBOCEAKNGYO--SDAKVYKTF 649
QY 609 LOWPYKXAGVBPBDBCKLFCARGSEPKVEAKYIDTLGAPETALCYRGQCYKAGCD 668
Db 650 VEWMPKTAGVLPADYCKLTCAPKGGYVVSPPKVTDTCEGRPFSNVSCVAGRCRYRCD 709
QY 669 HAVDSPRKLDKGVYCGKGNSCRKYSGLPFTNGYNDIVTIPAGATINIDVKQSRHPGVQ 728
Db 710 GIISSKIOYDKGVGGGNDSSCTKILGFNNKSKSYDTVAIPBEGATHIKVQEFAXDQT 769
QY 729 NDGNVLAIKTADGOYLNGNLALSAIEDILVKYTIILKSSIALTERL--QSRPLPEP 786
Db 770 RFPALALKKKTGTEYLINGKYMISTSEETIIDINGTVMNYSQMSHRDPLHGMSATKEI 829
QY 787 LTVQOLTPGGEVPEPK--VKYTFPVNDVDPSMOSKERATYNIQPLLAHOWLYGWS 843
Db 830 LIVQOLATD---PFKALGVRSFVPEPKKTYKQNVSVSHSGNKRGPHSTQLQWYTGFWL 885
QY 844 ECSSFCGAGMORATECRDPGSGOASATCNKALKPEADAPCE5QLC 888
Db 886 ACSRICDTGMHRTIVQCQDGNRKLAKGLLSGRPAFQCLLKCC 930

RESULT 11
US-09-122-126B-15
; Sequence 15, Application US/09122126B
; Patent No. 6451575
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909
; CURRENT APPLICATION NUMBER: US/09/122,126B
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 930
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-122-126B-15

```

```

Query Match 35.7%; Score 1734.5; DB 4; Length 930;
Best Local Similarity 39.8%; Pred. No. 2,8e-138;
Matches 378; Conservative 122; Mismatches 348; Indels 101; Gaps 20

QY LLLLLLLPLARGAPRRPA---GGQ---ASELVVPTRLG----- 49
    |||
Db LLLCAFRLLPAAVGFPATPAQDAKGPFPAAAAQPRRQGEVQERAEPPGHFHLAQR 67
    |||
QY 50 -----SAGELALHLISAFKGFVLRLAPDPSFLAPEKIERLGGSGRATG 94
    |||
Db 68 RRSKGLVQNIQDLYSGGKGVYLVAAGRRFLDLDRDSV-----GIAGFVPAGG 119
    |||
QY 95 E-----RGLRCFSGVYNGEPESLAVSICRGISGSEFLIDGEBFTIOPGAGSLQPPR 150
    |||
Db 120 TSAPMWRHSHCFRGVTDASPRSLAFLDGLGIDGFPFAYVHAYTLKPL-LRPPMAEEK 178
    |||
QY 151 LQRMGPAGCAPLPRGPEWEVETGEGORQ---RGDHQEDSEESQEEELAGASEPP--- 203
    |||

```

```

Db      179 GRVYGDSASAIL-----HYTRREGFSFEALPRASCEPASTPEAEHEPAHNSPSGRA 232
Qy      204 -----PLGATIS-----RKRVSEARFETELLVVDASAAFGALJONHIL 245
Db      233 ALASQILLDQALSPAGSGSPQJTWRRRRRSTISRARQVELLLVDASAAFLYGRBLQYTL 292
Qy      246 TMSVAARIYKHSPIKNSINLMVYKYLIVEDKMPSPVSDNGGLTLNFCNMQRRFNQPS 305
Db      293 TLASIRARLVSHASIENTHIRLAVKVVVLGDKSLSEVSKAAATTLNFCMCHQHNQLG 352
Qy      306 DRPEHYDTAILLRQNFQCGOGLCTLGVADIGTIDCPMKSCSVIEDGEQAHTLAHE 365
Db      353 DDHEEHDAALLFRREDLCGHS--CDPLGNADVGTICSPESCAVIEDDGLHAAFTAAHE 411
Qy      366 LGHTLSPMDHDSKCTRLFGPMGKHHTMAAPLFVIANOTLPMWPCSAWYLTBLDGGHDC 425
Db      412 IGHLLGSHSDSKCESTFSGTBDKRLMSSLTISDASKPMKCTSATITEFLDGHGNC 471
Qy      426 LLDAPGALLPLPTGLPGHMAIYOLDQOCROJFGSDPFHCPNTSADQVQAOLMCH--TDOAE 484
Db      472 LLDLPRKQILQIBELPFG--TYDATQOCNLTFGSEYVCP--GMDVARIKMACAVVRGQ 526
Qy      485 PLCHTKNGSLPMAAGTPCGPGLHSESGCLPEEVEERPKPVAVDGGMAFWMGEGCSRTCG 544
Db      527 MWCLTK--KLVAVESTPGCKRILCLQKQCVDKTKKKYSTSNGMWGSGWGQCSFSCG 584
Qy      545 GGVQFSHRECKDPEPONGARYCLGPRAYQSCHTBECPDDKSPFEQCEKKNANMYDM 604
Db      585 GGVQFAYHNCNPPARNNGRYCTGFRALYRCSLMPCPFNKSPFHEQCEANRGQ--SDA 643
Qy      605 DG--NLLOWVRYKAVGSPRDKCKLFCRARGSEFVFAKYIDGLCCPETLALCVRSQC 662
Db      644 KGAVTKFVMPVRYKAVLPAADVCKLTCRAKGGYVYVFPKXTLDGECRPSYNSVCVRGKC 703
Qy      663 VKAGCDHVVDSPRKLDKCGVCGGKNSCRKYSGLTPINNGYNDLYVTPAAGTINDVOR 722
Db      704 VRTGCDGIGSLQYDKCGVCGGDNSSCTKIVGTNNKSKGTTDVVRLEGHTHLKVRQF 763
Qy      723 SHPGVQNDGNVYALKTADQGYLNNMLNSAIEQDILVKGTLILKYSJATLERL--QSF 780
Db      764 KAKQDRTTAYALAKKNGEYLINKMWISSETIIDLNGVTMNSYSGSHRDPFLHGMGY 823
Qy      781 RPLPEPLTVQOLI--TVPGEVFPPEPKYKTYTFPVNDVDFSMQSSHERKATINIIGQLLAAOWLV 839
Db      824 SATKRIILLIVQILIADPTK--PLDVAYSFVPEKSGTSPKXNSVTSHGSKNGASTTSQPOWVT 881
Qy      840 GPMSECSSTCGAGWQRATYECDDPSGQASATCNKALKEDAKPCEBOLC 888
Db      882 GFWLACSRCTCDGHTTRIVQCQDGNRKLAKGSPQSRSAFQCLKKC 930

```

```

RESULT 12
US-09-634-286A-15
; Sequence 15, Application US/09634286A
; Patent No. 6521436
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909A
; CURRENT APPLICATION NUMBER: US/09/634,286A
; CURRENT FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-634-286A-15

Query Match      35.7%; Score 1734.5; DB 4; Length 930;
Best Local Similarity 39.8%; Pred. No. 2.8e-138;
Matches 378; Conservative 122; Mismatches 348; Indels 101; Gaps 20.

```

Fri May 7 12:18:01 2004

us-09-989-687-4\_1.ra1

Page 8

15 LLLLLLLPLAGAPAPAA---GGQ---ASELVPTRLPG----- 49  
8 LLLCAFLPLAAGVAPATPAODKAGPPTAAAAAQRORRQGEVEGRAPPPHPLAOR 67  
50 -----SAGELALHLSAFKGFVLRPLAPDPSFLAPEKTERLGGSGRATGG 94  
68 RREKGLVONIDOLYSGGKGVLYVAGGRFLDLERDSV-----GIAGFVPAAGG 119  
95 E---RGLRGCFPSGVNGBEPESLAVSICRGLSGFLDGBEFTIQPGAGGSLAQPHR 150  
120 TSAPWHRSHCFYRGVVDASFRSLAVFDLGGDGFAYKHAARYTLKPL-LGPPAEEBK 178  
151 LQRMGAPAGAPLRPGPEWEYETEGOROE---RGDHOEDSEESQEEAEAGASEPP--- 203  
179 GRVYGGDSARIL-----HYTREGSFELPRACCEFPAGTPEHNEHAPHSNSGRA 232  
204 -----PLGATS-----RTKRFVSAPFVETLLVADSMALFYGADLONHL 245  
233 ALASQLDDQALSPAGSGSPQTMWRRRRRSISRARQVELLVADSMALFYGRLQHYLL 292  
246 TLMSVARIYKHSIKXSINLMVYKYLIVDEKMGPEVSDNGGLTRFCNNQARFNPQS 305  
293 TLASIAIRLYSHASIEHILAVKVVYLDKXSLFVSQNAATLTKFCXKHQHNQ 352  
306 DRHEHYDTAILLRQFCGQEGELCDTLGVADITGTCDPNKSQSVTEDEGLQAHTLAHE 365  
353 DHEHYDAIILFREDLCGHS-CDTLGMADVGTICSPERSCAVIEDGLHAFTVAHE 411  
366 LGHVLSPHDSKRCCTFLPGMGNHMAFLFVHLNQTLPSPCSAMVLTLLDGDGHC 425  
412 IGHLLGSHDSKRCCTFPOSTEDKLMSSILTSIDSKMSKCTATITEFLDDHGHC 471  
426 LLDPAALPLPTGLPGRMALYOLDQCRQIFGDFHCPNTSAQDVCAQJMC-TDGAB 484  
472 LLDPRQOILGPBELPGO--TYDATQCCNLTFGEYSVCP---GMVCAALMAVVRQO 526  
485 PLCHTKNGLPMADGTPCGRHLCSBESCLPEEVEHPRPVDDGMAPMGECERTG 544  
527 MYCJTK--KLPAVGTGCGKRICLGKCVDKTKKTYTSSHONMGSWQCSBSCG 584  
545 GGVQFSHRECKDPKPPONGRXYCLGRRAKYOSCHTECPDPKSPREOCCEKYNAYYTM 604  
585 GGVQFAYHCKNPAFNNNGRXYCTGKRAIYSCSLMPCPPNGKSPRHQCEAKNGYQ-SDA 643  
605 DG--NLQWVRYKAVSPDRCKYCFARAGRSEKPYEAKYIDGTLCGPETLALCYRGCC 662  
644 KGVTFVEMVYKAVGVPADVCKLTCRAKGTGYVVSFKYTDGTECRPFYSNSVCVRGKC 703  
663 VKAGCDHVDSRKLDKGVCYGGKNSCRKYSGLTFPTYGYNDIVTIPAGATNIDVKOR 722  
704 VRTGDDGILGSKLDKDGCGVCGDSSCTKIVGTFNKSKGYTDVVAIPGATHTIKXRF 763  
723 SHPGVQNDGNIALKTADGOYLNGNTAISALBODILVKITLKYSGSIATLEL--QSF 780  
764 KAKODTRFAYLALAKKNGEYLINGKYMISTEITIIDINGVMYVSGSHRDDFLHMGY 823  
781 RPLRPLLYOLL--TPRGEVPRKXYTTFVNDVDFSMQSKERATNIIOLPLHAQVYL 839  
824 SATKILLVQILADPFR--PUDVAYSFVPEKSTPKVNSYTHSGSNKVGSHTSQPOWVT 881  
840 GDMSECSSTCGAGMORFTECDPSPGASATCNKALKEDEAKPESQJLC 888  
882 GFWLACSRCTDGTGHTRTVQCDDGRKLAKGCPLSQPSAFKQCLLKKC 930

RESULT 13  
US-09-369-364A-13  
Sequence 13, Application US/09369364A  
Patent No. 6391610  
GENERAL INFORMATION:  
APPLICANT: Apte, Suneel  
APPLICANT: Hurskainen, Tiina L.  
APPLICANT: Hirohata, Satoshi

TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/4007/10-30-00  
CURRENT APPLICATION NUMBER: US/09/369,364A  
CURRENT FILING DATE: 1999-08-06  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 13  
LENGTH: 1882  
TYPE: PRT  
ORGANISM: Homo sapiens ADAMTS-9  
FEATURE:  
NAME/KEY: MOD RES  
LOCATION: (468)  
OTHER INFORMATION: Xaa = C  
NAME/KEY: MOD RES  
LOCATION: (521)  
OTHER INFORMATION: Xaa = Y  
US-09-369-364A-13

Query Match 33.1%; Score 1608; DB 4; Length 1882;  
Best Local Similarity 37.5%; Pred. No. 4.5e-127; Indels 138; Gaps 24;  
Matches 339; Conservative 150; Mismatches 276;

58 LSAFGKGVFLRLADDSFLAPEFKTERLGGSG-----RATGGERGLRGCFPSGVNGBPE 112  
104 LSAFGQQLFVLTNAGFIAPLFTVTLTGTVNQTFYSEBAELKHCFYK----- 155  
113 SLAASVLCRGLSGFLDGBEFTIQPGAGGSLAQPH-----RLQRMGP----- 156  
156 -----RLC-----OYLRAHGRHPLRLNEHKNRHSQDKKTPARKGERINTL 198  
157 AG-ARPLRPGPEWEYETEGORORERGDHOEDSEESQEEAEAGASEPPPLGATSRKYF 215  
199 AGDVAAALSGLATEFAKYG-----NKTDNTEKTRH-----RTKRF 236  
216 VSEARFVETLLVADSMALFYGADLONHLTLMSVARIYKHSIKXSINLMVYKVLIVE 275  
237 LSPRFVFLVADRMVSYHGENLOHYILTLMSIVASTYDPISGLINIVINLVITH 296  
276 DEKMPPEVSDNGGLTRNFCNMORRFPNOSPDRHEHYDTAILLRQFCGQEGELCDTLGV 335  
297 NQDPPSISFNQATLTKFCQWHS-NSPG--IHHTAVLLRQDICRAHDKCDTLGL 352  
336 ADIGTCDPNKSQSVTEDEGLQAHTLAHEIGHVLSVPHDSKRCCTFLPGMGNHMAF 395  
353 AELGTCDPYRSCSISEDSGLSTAFTHAELGHVFNMHDDNNKC-KEEGVKSPOHVNAF 411  
396 LFVHLNQTLPMSPCAMVLTLLDGGHDCILLDAP-GAALPLPTGLPGRMALYOLDQCR 454  
412 TLFYTNPMWMSKSRKTYTEFLDTGCECLINEPESRPLPVQLPG--ILYNNKQXE 469  
455 QIFGDFHCPNTSAQDVCAQJMC-HTDGAELCHTKNGLPMADGTPCGRHLCSBSC 513  
470 LIFGSGQYCFYMWQ---CRLLMNNVNGVHKGCRTOH--PFWAGTECEPEKHCXQFC 524  
514 LBEVEYERKPYVDGMAFPMGMECSRTCCGGVQFHSRECKDPKPPONGRXYCLGRAPY 573  
525 VPK--EMDVPTDGSWGSPPGTSRCGGGIGITAIARECNRPRPKAGKCYVRBMKF 581  
574 QSCHECPDPKSPREOCCEKYNAYYTMNGNT--LQWVRYKAVSPDRCKYCFAR 631  
582 KSNTPPLCKKORDPFDECAHFDGHF--NINGLLPNVWVRYSGIILMKDRCKLFCRYVA 640  
632 GRSEFVFEAKYIDGTLCGPETLALCYRGQCYKACGDHVDSRKLDKGVCYGGKNSCR 691  
641 GNTAVYQLDRVIDGTPCGQDINDICVQGLCSQAGCDHVLNSKARDKCGVCGDSSCK 700  
692 KYSGSLPTNYGYNDIVTIPAGATNIDVKORSHPGVQNDGNIALKTADGOYLNGNTAI 751  
701 TVAGTNTYHIGNVYLRIPAGATNIDVQHSFSGETDDNYALSSSKGEFLNGNFV 760  
752 SAIEDILVKITLKYSGSIATLERQSFRLPEPLTYOLLVVPGEVFPKXYTFEYFN 811



Fri May 7 12:18:01 2004

us-09-989-687-4\_1.rat

Page 9

```
Db 761 TNAKREIRIGNAVESGETAVERINSTDRIEGELLLOVLSV-GKLYNPVRSFNPI 819
Qy 812 D-----VDFSMQSKERATT-----NTIQPL- 832
Db 820 EDKPOFYNNSHGPMWQACSKPCQGERKRLVCTRESQDLTVSDQRCDLRPPGHITEPCG 879
Qy 833 --LHAQWVLGDMSECGSTGAGWQKRTVECRDPSGQASAT-----CNKALKEPADKPC 883
Db 880 TGCDDLHWVAHSRSECSAQCGLGRTLDYCAKYSRLDGTEKVDGFCSSHPKPSNRXC 939
Qy 884 ESQ 886
Db 940 SGE 942

RESULT 14
US-09-369-364A-15
; Sequence 15, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Aptec, Suneel
; APPLICANT: Harkkainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ. ID NOS.: 31
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 15
; LENGTH: 874
; TYPE: PRT
; ORGANISM: Mus musculus ADAMTS-9
US-09-369-364A-15

Query Match 32.2%; Score 1562; DB 4; Length 874;
Best Local Similarity 37.3%; Pred. No. 1,1e-123;
Matches 322; Conservative 152; Mismatches 266; Indels 124; Gaps 22;

Qy 115 AAVSLCGSLGSLFDLDEEFTIOP-----QAGGSLAOPHRLQRMGPAGAPRLPRGPPM 168
Db 3 AVISLCGGMGTFRSHGDTFIEPLQSVDBQDEBEQNKPHIYRST-----PGR 53
Qy 169 EVETGE--GQROERGHQSDS-----EEESQEEKAGASEPPPLGAT----- 209
Db 54 EPSTGKACATSELSKNSKDKIRMRKRRKNSIADVDALKSGIATKVLGSGYNQTN 113
Qy 210 -----SPTKRVSAAREVELTLLVADASMAAFYGADLQNHITLMSVAA-YYKHPSTK 261
Db 114 NTRDRMNRKRTKRLSLPRFEVNVVADHRMVLTHGNLQHTLTLMSIYASISYKSSIG 173
Qy 262 NSINIMVVKVLIIVDEKMGPEVSDNGGLTLRNFQWRRRNPQSDRPHFY-----DTA 315
Db 174 NLINIVLVNVLVINEGEGPYINFMNQTLKNFCQWQ-----HSKNYLGSIQHDYA 224
Qy 316 ILLTRQNFQCGEGICDTLGVADITICDPNKSQSVIIDEGLQAHLTAHLGLVLSMPH 375
Db 225 VLVTRREDICRAQDCKDTLGLAELGTICDPYRSCSISDSGLSTFTIHLGLHGVFNPHD 284
Qy 376 DSKCCTRLFGPMGKHVWAPLFLVHINQTLFWSPCSANYLTLLDGGHCDCLDAPGA-AL 434
Db 285 DSNKC-KEEGVKSQHVWAPLFLVHINQTLFWSPCSANYLTLLDGGHCDCLDAPGA-AL 434
Qy 435 PLFTGLPGRVALYOLDQCCROIIFQPDRLHCPNTSAQDVCAQWNC-HTDGAPLCHTKNGS 493
Db 344 PLPQQLPQ--LLYVNNKQCELIIFGSGQVCEYMQO-----CRLMNNVDGAHKQKTOH-- 396
Qy 494 LPMADGTCCGPGHLCSESGCLPEEVEVERPKPVVDGVGNAFPMGPWGCSCRTGGGVQFSHRE 553
Db 397 TFMADGTCEBCKKCKKCFGVK--EVEGPAIDGSGWGMHFGTCSRTGGGKTAIRE 453
Qy 554 CKDEPONGARYCIARRAKYQSCHTBECPPDGKSFREQCEKTAANYTMDGL--LQW 611
```

```
Db 454 CNRPEPNQGVKVCGRMRKFSNTEPCKMKQKDFREECQAFHDPKHF-NINGLLPSVRW 512
Qy 612 VPKYAGVSPDRCLTFCRAGRSSEFKVFEAKVLDGTLGSETLLACRGQCVVAGDHY 671
Db 513 FPKYSGLMDRCCLTFCRVAGNTAYQLRDRVIDGTPCGQDTNDICVQGLCRQNGCHIL 572
Qy 672 DSPKRLKCGVCGGKSGKRSKVSGLTPVNYGYNDIYTPAGATNIDVKQSRHPGVQNDG 731
Db 573 NSKTRKCKGICGGSDNSCKTVAQTFTVYGVNTVIRIAGATSIDVRQHSFGKSEDD 632
Qy 732 NYLAKTADQVLYNLNLAISALEQDILVKTGLIKSGSLATLERLQSPRLPPLTVOL 791
Db 633 NYLALNSKGEFFLNGDFVYVMSKREVRVSAVLEYSQDNVCERLNCTDRIBELLQV 692
Qy 792 LTVGEVPPKVKYTFVVPNDVDFSMQ----- 818
Db 693 LSV-GKLYNPVRSFNPIE-DKPOFYNNSHGPMWQACSKPCQGERKRLVCTRESQDL 750
Qy 819 SKKERATTNITQ--PLIHA-----QWLGDMSECGSTGAGWQKRTVECRDPSGQASA 869
Db 751 TVSDQRCDLRFPQBPVTEACGTDCDLRWVHVASRSECSAQCGLGRTLDYCAKYSRMDGK 810
Qy 870 T-----CNKALKEPADKPCESQ 886
Db 811 TEKVDSEFSCQPRPSNOEKCSGE 834
```

```
RESULT 15
US-09-800-729-155
; Sequence 155, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ. ID NOS.: 217
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 155
; LENGTH: 2165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-155

Query Match 30.4%; Score 1476; DB 4; Length 2165;
Best Local Similarity 34.1%; Pred. No. 8,9e-116;
Matches 327; Conservative 150; Mismatches 319; Indels 164; Gaps 29;
```

```
Qy 46 RLPGSAGEIALHLISAFGKGVRLAPDPSFL-----APEFKIRLDGS 88
Db 96 RLQVARDQ-----GHACHRLRLSDDAVYIVLHRMNOIPDSHNKSVPHRSNSNPAW 148
Qy 89 GKAATGEGRLG-----CFPSGTVNGEPELSAAVSLC---RLSGSFLDGEFTIQ 137
Db 149 VLYLDSSEBVRGMSKRTPDCTIRAHVAGVHGH-SIVLDCSEBDLYGMALLPSGIHIVE 207
Qy 138 P-----QAGGSLAOPHRLQRMGPAGAPRLPRGPEWEVETGEGQROERGDH-QEDSEE 189
Db 208 PIIISGTEHDASHRHRLVAKFPDPMHFKP-----DHLNSTSVN 248
Qy 190 ESQEEBAASBPPLGATSRTKFY-SEARFVETLLVADASMAAFYGADLQNHITLM 248
Db 249 ETEITVAITQDMEDVIERKARSRRAANSMDHYEVVLYAADTKMEYHGRSLSEDVLTIF 308
Qy 249 SVAAIRYKPSIKNSINIMVVKVLIIVDEKMGPEVSDNGGLTLRNFQWRRRNPQSDRPH 308
Db 309 STVASIYRQSIKRAINVVVVLIVLKTENAGPRITONAAQDTLDFCRWQOYNDPDSS 368
```

Fri May 7 12:18:01 2004

us-09-989-687-4\_1.rtf

Page 10

QY 309 PEHYDTAILITRONFCQOEGLCPTLGVADIGTICDPNKSCSVIEDEGLQAAHTLAHELGH 368  
DB 369 VQHHDDVALITLRKIDICRSQKCDTLGLAELGTMCOKSCALIEDNGLSMAFTIAHELGH 428  
QY 369 VLSMPHDDSKPCSTRLPQPMGK-----HVMAPLFVHLNQTLPMSPCSA 411  
DB 429 VFSIPHDDEKCS-TYMPVNVKCFQSTKFDKTOFQNNFHIMAPLTLEYTHPMSWSPCSA 487  
QY 412 MYLTLLDGGHG--DCLLDAPGAAL----PLPTGLPGRMALYQLDQOCROIFEGDFRHP 465  
DB 488 GMLERFLENNRGOTQCLFDQVEBRYEDVFVRDEPKK--YDAHQCKVFQFASLCP 545  
QY 466 NTSADVCACQIMCHT-DGAEPLCHTKNGSLPMADGTPCGPH--LCSGSL--PEEV 519  
DB 546 ---YMPCTRRLMCAITFGSQWGCRTGH--MPNADGTPCDESRSMFCHHGACVRLAPESLT 600  
QY 520 ERFPVVDGMAFPMGEGSCRTGGVQSHRECDPEPONGRYCLGRAYQSCHE 579  
DB 601 K-----IDGQWDRSWGECSTGCGVQKGLRDCSPKPRNGKTCVQGRERYSCNTQ 655  
QY 580 ECPDGKSFRCQCEKTN--AYNYTDMGNLQWPKYAGVSPRDCKLFCRARGSEPK 637  
DB 656 ECFMDTQPYREVQCFSENNKDIGIQVASTNTHWPKYANVAPNERCKLYCRLSGSAFY 715  
QY 638 VFEAKVIDGTLGPEETLACVRCQCVACDHYVDSPRKLDKCGVCGKGNCKRYSGSL 697  
DB 716 LLRDKVVDGTPCDRNGDDICVAGACMPAGCDHQLHSTLRDKCGVCGSDSCVKYKGT 775  
QY 698 TPT-NGVNDIVTIPAGATNIDVQKSHPGVQNDGNLAKTADGOYLNGNLASIEQ 756  
DB 776 NEQGTFGYNEWMKIPASANDIRKQKXNNKEDDNYLSRAANGELFNGHFOVSLARQ 835  
QY 757 DILVGTILKXYSIATLERLOSFRPLPEPLTVQLTVPGEVPPKVKYTF--VPNDY 813  
DB 836 QIAFQDVTLEYSGSDAIIERINGTPIRSDIYVHLSVGSN--PPDISYEMTAAPNAV 893  
QY 814 -----DFSM-QSKERATNIIOL----- 832  
DB 894 IRPISALYIMRVITDTECDRACRQSQSQKLMCLDWSHTRQSHDRVCQNVLPKQATRM 953  
QY 833 ---LHAQVVLGDMSECSSTGAGMORRTVECRDPSG--QASAT--CNKALKEBDAKPC 883  
DB 954 CNIDGSTWITEDEVSSGACXGSGQKQKQVSCVMMEGBRQTPASEHLCDRNSKEPDIASC 1013

Search completed: May 7, 2004, 11:52:17  
Job time : 26 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:57:22; Search time 23 Seconds  
(without alignments)  
1997.701 Million cell updates/sec

Title: US-09-989-687-4  
Perfect score: 890  
Sequence: 1 MFPAAPAPRWLPFLILL.....CNKALKEPDAPKESQGLPL 890

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 389414 segs, 51625971 residues

Word size: 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

Issued Patents AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A-COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5E-COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A-COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B-COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCUS-COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfill1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	217	24.4	245	US-09-369-364A-11	Sequence 11, Appl
2	35	3.9	905	US-09-369-364A-9	Sequence 9, Appl
3	30	3.4	481	US-09-130-491-8	Sequence 8, Appl
4	13	1.5	2165	US-09-800-729-155	Sequence 155, App
5	12	1.3	438	US-09-963-791-22	Sequence 22, Appl
6	12	1.3	551	US-09-130-491-16	Sequence 16, Appl
7	12	1.3	589	US-09-963-791-12	Sequence 12, Appl
8	12	1.3	727	US-09-445-023A-12	Sequence 12, Appl
9	12	1.3	757	US-09-963-791-24	Sequence 24, Appl
10	12	1.3	837	US-09-122-126B-2	Sequence 2, Appl
11	12	1.3	837	US-09-634-286A-2	Sequence 2, Appl
12	12	1.3	859	US-09-369-364A-5	Sequence 5, Appl
13	12	1.3	908	US-09-963-791-2	Sequence 2, Appl
14	11	1.2	263	US-08-300-903A-2	Sequence 2, Appl
15	11	1.2	263	US-08-988-197-2	Sequence 2, Appl
16	11	1.2	655	US-08-148-910-12	Sequence 12, Appl
17	11	1.2	655	US-08-448-937A-12	Sequence 12, Appl
18	10	1.1	58	US-09-800-729-168	Sequence 168, App
19	10	1.1	205	US-09-800-729-149	Sequence 149, App
20	10	1.1	207	US-09-800-729-113	Sequence 113, App
21	10	1.1	208	US-09-800-729-151	Sequence 151, App
22	10	1.1	231	US-08-220-379B-7	Sequence 7, Appl
23	10	1.1	231	US-08-243-545-2	Sequence 2, Appl
24	10	1.1	231	US-08-993-962-2	Sequence 2, Appl
25	10	1.1	231	US-09-160-841-2	Sequence 2, Appl
26	10	1.1	231	US-08-669-692-2	Sequence 2, Appl
27	10	1.1	231	US-08-444-626-2	Sequence 2, Appl

28	10	1.1	231	5	PCT-US94-05365-2	Sequence 2, Appl
29	10	1.1	231	5	PCT-US95-03866-6	Sequence 6, Appl
30	10	1.1	467	4	US-09-148-545-134	Sequence 134, App
31	10	1.1	467	4	US-09-907-794A-195	Sequence 195, App
32	10	1.1	467	4	US-09-905-125A-195	Sequence 195, App
33	10	1.1	467	4	US-09-902-775A-195	Sequence 195, App
34	10	1.1	480	2	US-08-828-488-8	Sequence 8, Appl
35	10	1.1	480	4	US-09-239-689A-8	Sequence 8, Appl
36	10	1.1	480	4	US-09-702-705-336	Sequence 336, App
37	10	1.1	480	4	US-09-736-457-336	Sequence 336, App
38	10	1.1	480	4	US-09-614-124B-336	Sequence 336, App
39	10	1.1	480	4	US-09-671-325-336	Sequence 336, App
40	10	1.1	480	4	US-09-589-184-336	Sequence 336, App
41	10	1.1	492	1	US-07-794-393-4	Sequence 4, Appl
42	10	1.1	492	1	US-08-001-711-4	Sequence 14, Appl
43	10	1.1	514	4	US-09-800-729-124	Sequence 13, Appl
44	10	1.1	608	4	US-09-130-491-13	Sequence 13, Appl
45	10	1.1	633	4	US-09-919-060-13	Sequence 13, Appl

#### ALIGNMENTS

RESULT 1  
US-09-369-364A-11  
Sequence 11, Application US/09369364A  
Patent No. 6391610  
GENERAL INFORMATION:  
APPLICANT: Apte, Suneel  
APPLICANT: Hurskainen, Tiina L.  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/4007/10-30-00  
CURRENT APPLICATION NUMBER: US/09/369,364A  
CURRENT FILING DATE: 1999-08-06  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 11  
LENGTH: 245  
TYPE: PRT  
ORGANISM: Homo sapiens ADAMTS-8  
US-09-369-364A-11

Query Match 24.4%; Score 217; DB 4; Length 245;  
Best Local Similarity 100.0%; Pred. No. 1.8e-203;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 196 AEGASEPPPIGATSRTRGVSEARFVETLLVADASNAAFGADLQNHILTLMSVAARIY 255  
DB 2 AEGASEPPPIGATSRTRGVSEARFVETLLVADASNAAFGADLQNHILTLMSVAARIY 61  
QY 256 KHPSTKNSINIMVAVLIVDEKMGPEVSDNGGTLIRFCNWMORFNPSPDRHPHYDTA 315  
DB 62 KHPSTKNSINIMVAVLIVDEKMGPEVSDNGGTLIRFCNWMORFNPSPDRHPHYDTA 121  
QY 316 ILLTRNQFCGEGGLCDTLGVADIGTICDPNKSQSVIEDEGIQAHTLAHEIGHVLSMHPD 375  
DB 122 ILLTRNQFCGEGGLCDTLGVADIGTICDPNKSQSVIEDEGIQAHTLAHEIGHVLSMHPD 181  
QY 376 DSKPCTRLFGPMGKHVNAPLFVHNLQTLPMSPCSAM 412  
DB 182 DSKPCTRLFGPMGKHVNAPLFVHNLQTLPMSPCSAM 218  
RESULT 2  
US-09-369-364A-9  
Sequence 9, Application US/09369364A  
Patent No. 6391610  
GENERAL INFORMATION:  
APPLICANT: Apte, Suneel  
APPLICANT: Hurskainen, Tiina L.  
APPLICANT: Hirokata, Satoshi  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases

FILE REFERENCE: 26473/4007/10-30-00  
CURRENT APPLICATION NUMBER: US/09/369,364A  
CURRENT FILING DATE: 1999-08-06  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 9  
LENGTH: 905  
TYPE: PRT  
ORGANISM: Mus musculus ADAMTS-8  
US-09-369-364A-9

Query Match 3.9%; Score 35; DB 4; Length 905;  
Best Local Similarity 100.0%; Pred. No. 2.9e-25;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 GVSFDRCKLFCRARGSEFFVFEAKYIDGTLGSP 651  
DB 632 GVSFDRCKLFCRARGSEFFVFEAKYIDGTLGSP 666

RESULT 3  
US-09-130-491-8  
Sequence 8, Application US/09130491  
Patent No. 6416974  
GENERAL INFORMATION:  
APPLICANT: Holtzman, Douglas A.  
APPLICANT: Goodheart, Andrew D.J.  
TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83  
FILE REFERENCE: 09404/041001  
CURRENT APPLICATION NUMBER: US/09/130,491  
CURRENT FILING DATE: 1998-08-07  
EARLIER APPLICATION NUMBER: US 60/058,108  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: US 60/054,961  
EARLIER FILING DATE: 1997-08-06  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 8  
LENGTH: 481  
TYPE: PRT  
ORGANISM: Rattus rattus  
US-09-130-491-8

Query Match 3.4%; Score 30; DB 4; Length 481;  
Best Local Similarity 100.0%; Pred. No. 1.2e-20;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 702 YGYNDIVTIPAGATNIDVKQRSHRPQVNDG 731  
DB 293 YGYNDIVTIPAGATNIDVKQRSHRPQVNDG 322

RESULT 4  
US-09-800-729-155  
Sequence 155, Application US/09800729  
Patent No. 6605582  
GENERAL INFORMATION:  
APPLICANT: Ni et al.  
TITLE OF INVENTION: 32 Human secreted proteins  
FILE REFERENCE: P2044P1  
CURRENT APPLICATION NUMBER: US/09/800,729  
CURRENT FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: PCT/US00/26013  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: 60/155,709  
PRIOR FILING DATE: 1999-09-24  
NUMBER OF SEQ ID NOS: 217  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 155  
LENGTH: 2165  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-800-729-155

Query Match 1.5%; Score 13; DB 4; Length 2165;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTGCGGV 548  
DB 612 WGECSRTGCGGV 624

RESULT 5  
US-09-963-791-22  
Sequence 22, Application US/09963791  
Patent No. 6649399  
GENERAL INFORMATION:  
APPLICANT: Donoho, Gregory  
APPLICANT: Turner, C. Alexander Jr.  
APPLICANT: Friedrich, Glenn  
APPLICANT: Scoville, John  
APPLICANT: Zambrowicz, Brian  
APPLICANT: Sands, Arthur T.  
TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the S  
FILE REFERENCE: LEX-0105-USA  
CURRENT APPLICATION NUMBER: US/09/963,791  
CURRENT FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: US 60/169,769  
PRIOR FILING DATE: 1999-12-09  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 22  
LENGTH: 438  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-963-791-22

Query Match 1.3%; Score 12; DB 4; Length 438;  
Best Local Similarity 100.0%; Pred. No. 0.0043;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTGCGGV 547  
DB 416 WGECSRTGCGGV 427

RESULT 6  
US-09-130-491-16  
Sequence 16, Application US/09130491  
Patent No. 6416974  
GENERAL INFORMATION:  
APPLICANT: Holtzman, Douglas A.  
APPLICANT: Goodheart, Andrew D.J.  
TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83  
FILE REFERENCE: 09404/041001  
CURRENT APPLICATION NUMBER: US/09/130,491  
CURRENT FILING DATE: 1998-08-07  
EARLIER APPLICATION NUMBER: US 60/058,108  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: US 60/054,961  
EARLIER FILING DATE: 1997-08-06  
NUMBER OF SEQ ID NOS: 16  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 16  
LENGTH: 551  
TYPE: PRT  
ORGANISM: Rattus rattus  
US-09-130-491-16

Query Match 1.3%; Score 12; DB 4; Length 551;  
Best Local Similarity 100.0%; Pred. No. 0.0053;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 706 DYTIPAGATNI 717  
|||||

DB 324 DIVTIPAGATNI 335

## RESULT 7

```
US-09-963-791-12
; Sequence 12, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedlich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399e1 Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-12
```

## Query Match

Best Local Similarity 1.3%; Score 12; DB 4; Length 589;  
Matches 12; Conservativity 0; Pred. No. 0.0056;  
Matches 12; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGCSCRTGGGV 547

DB 567 WGCSCRTGGGV 578

## RESULT 8

```
US-09-445-023A-12
; Sequence 12, Application US/09445023A
; Patent No. 6565858
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioke, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE REFERENCE: 057092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; CURRENT FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-445-023A-12
```

## Query Match

Best Local Similarity 1.3%; Score 12; DB 4; Length 727;  
Matches 12; Conservativity 0; Pred. No. 0.0068;  
Matches 12; Mismatches 0; Indels 0; Gaps 0;

QY 706 DIVTIPAGATNI 717

DB 500 DIVTIPAGATNI 511

## RESULT 9

```
US-09-963-791-24
; Sequence 24, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedlich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399e1 Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 757
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-24
```

## Query Match

Best Local Similarity 1.3%; Score 12; DB 4; Length 757;  
Matches 12; Conservativity 0; Pred. No. 0.007;  
Matches 12; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGCSCRTGGGV 547

DB 416 WGCSCRTGGGV 427

## RESULT 10

```
US-09-122-126B-2
; Sequence 2, Application US/09122126B
; Patent No. 6451575
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909
; CURRENT APPLICATION NUMBER: US/09/122,126B
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-122-126B-2
```

## Query Match

Best Local Similarity 1.3%; Score 12; DB 4; Length 837;  
Matches 12; Conservativity 0; Pred. No. 0.0077;  
Matches 12; Mismatches 0; Indels 0; Gaps 0;

QY 539 CSRTGGGVQPS 550

DB 532 CSRTGGGVQPS 543

## RESULT 11

```
US-09-634-286A-2
; Sequence 2, Application US/09634286A
; Patent No. 6521436
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909A
; CURRENT APPLICATION NUMBER: US/09/634,286A
; CURRENT FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.0
```

```

; SEQ ID NO 2
; LENGTH: 837
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-634-286A-2

Query Match          1.3%; Score 12; DB 4; Length 837;
Best Local Similarity 100.0%; Pred. No. 0.0077;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 CSRTCGGCGV 550
DB 532 CSRTCGGCGV 543

RESULT 12
US-09-369-364A-5
; Sequence 5, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apple, Suneeel
; APPLICANT: Hurskainen, Tina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 5
; LENGTH: 859
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-6
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (450)
; OTHER INFORMATION: Xaa = L
US-09-369-364A-5

Query Match          1.3%; Score 12; DB 4; Length 859;
Best Local Similarity 100.0%; Pred. No. 0.0079;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGV 547
DB 519 WGECSRTCGGV 530

RESULT 13
US-09-963-791-2
; Sequence 2, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399 Human Proteases and Polynucleotides Encoding the SA
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 908
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-2
```

```

Query Match          1.3%; Score 12; DB 4; Length 908;
Best Local Similarity 100.0%; Pred. No. 0.0083;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGV 547
DB 567 WGECSRTCGGV 578

RESULT 14
US-08-300-903A-2
; Sequence 2, Application US/08300903A
; Patent No. 5581630
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M
; APPLICANT: Gail, Judith G
; TITLE OF INVENTION: Interleukin-15 Receptors
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple Operating System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/300,903A
; FILING DATE: 06-SEPTEMBER-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/236,919
; FILING DATE: 06-MAY-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2822-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; TELEFAX: 206-233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 263 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-300-903A-2

Query Match          1.2%; Score 11; DB 1; Length 263;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLLPL 24
DB 18 LLLLLLLLLPL 28

RESULT 15
US-08-988-197-2
; Sequence 2, Application US/08988197
; Patent No. 6548065
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M
; APPLICANT: Gail, Judith G
; TITLE OF INVENTION: Interleukin-15 Receptors
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
```

Fri May 7 12:17:59 2004

us-09-989-687-4.ra1

Page 5

STREET: 51 University Street  
CITY: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Macintosh  
OPERATING SYSTEM: Apple Operating System 7.1  
SOFTWARE: Microsoft Word for Apple, Version 5.1a  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/988,197  
FILING DATE:  
CLASSIFICATION: 121097  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/300,903  
FILING DATE: 06-SEPTEMBER-1984  
APPLICATION NUMBER: USN 08/236,919  
FILING DATE: 06-MAY-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2822-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-587-0430  
TELEFAX: 206-233-0644  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 263 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-988-197-2

Query Match 1.2%; Score 11; DB 4; Length 263;  
Best Local Similarity 100.0%; Pred. No. 0.026;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLLL 24  
Db 18 LLLLLLLLLL 28

Search completed: May 7, 2004, 12:01:23  
Job time : 24 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:51:46 ; Search time 65 Seconds  
(without alignments)  
3868.727 Million cell updates/sec

Title: US-09-989-687-4

Perfect score: 890  
Sequence: 1 MFPAAPAPRMVLFLLLL.....CNKALKPEDAKPCESQLCPL 890

Scoring table: OLIGO  
Gapop 60.0 , Gapept 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: Geneseqp1980s:\*\n2: Geneseqp1990s:\*\n3: Geneseqp2000s:\*\n4: Geneseqp2001s:\*\n5: Geneseqp2002s:\*\n6: Geneseqp2003as:\*\n7: Geneseqp2003bs:\*\n8: Geneseqp2004s:\*

Prod. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	890	100.0	890	2	AAV49502 Human MET
2	890	100.0	890	4	AAV49502 Human MET
3	675	75.8	890	6	ABP96306 Human ADA
4	649	72.9	889	4	AAV4946 Human ADA
5	362	40.7	680	3	AAV21251 Human met
6	263	29.6	364	4	AAV09710 Human gen
7	263	29.6	364	7	ADCC2088 Human pro
8	217	24.4	245	4	AAV72285 Human ADA
9	35	3.9	905	4	AAV72284 Human ADA
10	30	3.4	481	2	AAV04145 Rat Tango
11	24	2.7	98	3	AAV21262 Rat meta
12	13	1.5	2150	3	AAV53898 Human aci
13	13	1.5	2150	3	AAV53898 Human aci
14	12	1.3	18	1	AAV81021 Sequence
15	12	1.3	19	4	AAV81021 Sequence
16	12	1.3	20	4	AAV81021 Sequence
17	12	1.3	54	3	AAV08136 Human aci
18	12	1.3	438	3	AAV08136 Human aci
19	12	1.3	575	7	ADV85489 Human aci
20	12	1.3	589	4	AAV03577 Human aci
21	12	1.3	625	5	AAV48394 Rat aggre
22	12	1.3	757	4	AAV03583 Human pro
23	12	1.3	837	2	AAV75425 Human aci
24	12	1.3	837	3	AAV98429 Human aci
25	12	1.3	837	4	AAV78228 Human aci

#### ALIGNMENTS

RESULT 1	AAV49502	standard; protein: 890 AA.
ID	AAV49502	
AC	AAV49502	
XX	10-JAN-2000	(first entry)
DT	10-JAN-2000	
DE	Human METH2 protein.	
XX	Human, METH1, METH2; anti-angiogenic; metalloprotease thrombospondin;	
XX	cancer; diagnosis; hyperproliferative disorder; autoimmune disease;	
XX	angiogenesis inhibitor; abnormal wound healing; inflammation;	
XX	rheumatoid arthritis; psoriasis; endometrial bleeding disorder;	
XX	diabetic retinopathy; macula degeneration; haemangioma; detection;	
XX	arterial-venous malformation; immune deficiency.	
OS	Homo sapiens.	
XX	MO9937660-A1.	
XX	29-JUL-1999.	
XX	22-JAN-1999;	99WC-US001313.
XX	23-JAN-1998;	98US-0072298P.
XX	28-AUG-1998;	98US-0098539P.
XX	(IRUE/) IRUELA-ARISPE L.	
XX	(HAST/) HASTINGS G A.	
XX	(RUBS/) RUBEN S M.	
XX	IrueLA-Arispe L, Hastings GA, Ruben SM,	
XX	WPI, 1999-590684/50.	
XX	DR N-PSDB; AA232001.	
XX	New isolated metalloprotease thrombospondin polypeptides, useful for	
XX	treating hyperproliferative disorders, cancers or autoimmune disorders.	
XX	Claim 10; Fig 2; 457p; English.	
XX	AA232000 and AA232001 encode, and AAV49501 and AAV49502 represent, human	
XX	metalloprotease thrombospondin (METH) proteins METH1 and METH2	
XX	respectively. METH1 and METH2 have been found to be potent inhibitors of	
XX	angiogenesis both in vitro and in vivo. They can be used for treating	
XX	cancer and other disorders related to angiogenesis including abnormal	
XX	wound healing, inflammation, rheumatoid arthritis, psoriasis, endometrial	





CC neovascularisation, telangiectasia, haemophilic joints, angiodiroma,  
 CC fibromuscular dysplasia, wound granulation, Crohn's disease or  
 CC atherosclerosis. METH2 can also be used in birth control. METH2 can also  
 CC be used in diagnostic methods for the prognosis of cancer

XX Sequence 890 AA;

Query Match 100.0%; Score 890; DB 4; Length 890;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPPAPAPRMLPFLILLILLILLPLARGAPAPAGQASSELVPTPLPGSAGELALHLSA 60  
 DB 1 MPPAPAPRMLPFLILLILLILLPLARGAPAPAGQASSELVPTPLPGSAGELALHLSA 60  
 QY 61 FKGFTLRLAPDPSFLAPEFKTERLGGSGRATGGERGLRGCFSGTVNGEPESLAAYSLC 120  
 DB 61 FKGFTLRLAPDPSFLAPEFKTERLGGSGRATGGERGLRGCFSGTVNGEPESLAAYSLC 120  
 QY 121 RGLSGSFLDGEFTIQPOGAGSLAQPHRLQWGPAGARPLPRGEWEVETGEGQORER 180  
 DB 121 RGLSGSFLDGEFTIQPOGAGSLAQPHRLQWGPAGARPLPRGEWEVETGEGQORER 180  
 QY 121 RGLSGSFLDGEFTIQPOGAGSLAQPHRLQWGPAGARPLPRGEWEVETGEGQORER 180  
 DB 121 RGLSGSFLDGEFTIQPOGAGSLAQPHRLQWGPAGARPLPRGEWEVETGEGQORER 180  
 QY 181 GDHQBDESEESQEEAEAGASEPPPLGATSRTRFVSEARFVETLLVADASAAAYGADL 240  
 DB 181 GDHQBDESEESQEEAEAGASEPPPLGATSRTRFVSEARFVETLLVADASAAAYGADL 240  
 QY 241 QNHILLTMSVARIYHPSTIKSINLMVYKVLVEDEKMGPEVSDNGGLTLNFCNQMR 300  
 DB 241 QNHILLTMSVARIYHPSTIKSINLMVYKVLVEDEKMGPEVSDNGGLTLNFCNQMR 300  
 QY 301 FNOPSDRPHRYDTAILTRQNFQCGEGLCDITGVADITICDPNKSCEVIDEGLOAH 360  
 DB 301 FNOPSDRPHRYDTAILTRQNFQCGEGLCDITGVADITICDPNKSCEVIDEGLOAH 360  
 QY 361 TLAEHGHVLSMHDSPKCTRLFGPMGKHWAAPFVLNQTLPSPCSAMYLIELDGG 420  
 DB 361 TLAEHGHVLSMHDSPKCTRLFGPMGKHWAAPFVLNQTLPSPCSAMYLIELDGG 420  
 QY 361 TLAEHGHVLSMHDSPKCTRLFGPMGKHWAAPFVLNQTLPSPCSAMYLIELDGG 420  
 DB 361 TLAEHGHVLSMHDSPKCTRLFGPMGKHWAAPFVLNQTLPSPCSAMYLIELDGG 420  
 QY 421 GHGDCULDAAGALPLPTGLPGHMAIYOLDQOCROIFGDFRHCENTSAQVCAOIWCHT 480  
 DB 421 GHGDCULDAAGALPLPTGLPGHMAIYOLDQOCROIFGDFRHCENTSAQVCAOIWCHT 480  
 QY 481 DGAEPLCHTNGSLPMWADGTPCGPHLCSEGLPEBEVERKPVVDGMAHPMGEGCS 540  
 DB 481 DGAEPLCHTNGSLPMWADGTPCGPHLCSEGLPEBEVERKPVVDGMAHPMGEGCS 540  
 QY 541 RTGGGVQFSHRECKDEPONGRVCIGRPAKYQSCHTBECPPDGKSPBQCEKXNAYN 600  
 DB 541 RTGGGVQFSHRECKDEPONGRVCIGRPAKYQSCHTBECPPDGKSPBQCEKXNAYN 600  
 QY 601 YTDMDGNLLQWPKYAGVSPDRCKLFCRARGSSEFVFAKYIDGLTLCPEFTLALCVAG 660  
 DB 601 YTDMDGNLLQWPKYAGVSPDRCKLFCRARGSSEFVFAKYIDGLTLCPEFTLALCVAG 660  
 QY 661 QCVNAGCDHVVDSPKLDKCGVCGKSGSKRYSGSLTPNYGNDIVITIPAGATNIDYK 720  
 DB 661 QCVNAGCDHVVDSPKLDKCGVCGKSGSKRYSGSLTPNYGNDIVITIPAGATNIDYK 720  
 QY 721 QRSHPGVQNDGNVLAATKADQVLLNGNLISAIBODILVKGITLTKSSGATLERLQSF 780  
 DB 721 QRSHPGVQNDGNVLAATKADQVLLNGNLISAIBODILVKGITLTKSSGATLERLQSF 780  
 QY 781 RPLREPLTVOLLTPGVEFPKXYTFFVNDVDFNSQSSKERATNIIQLPLHAQVNLG 840  
 DB 781 RPLREPLTVOLLTPGVEFPKXYTFFVNDVDFNSQSSKERATNIIQLPLHAQVNLG 840  
 QY 841 DMSGCSSTCGAGWQRRTVECRDPGQASATCNKALKEDAPCESQICPL 890  
 DB 841 DMSGCSSTCGAGWQRRTVECRDPGQASATCNKALKEDAPCESQICPL 890

RESULT 3

ABP96306  
 ID ABP96306 standard; protein; 890 AA.

XX ABP96306;

AC 20-MAY-2003 (first entry)

XX Human ADAMTS8 protein.

XX Humanised baculovirus; cytosolic; gene therapy; baculovirus; cancer;  
 KM prostate cancer; chromosome 11.

OS Homo sapiens.

PN W02003016540-A2.

PD 27-FEB-2003.

PF 15-AUG-2002; 2002MO-GB003791.

PR 15-AUG-2001; 2001GB-00019852.

PA (UYO-) UNIV YORK.

PI Maitland N;

DR WPI; 2003-268336/26.

PT New baculovirus having a modified genome encoding a therapeutic agent,  
 PT useful in the manufacture of a medicament for the treatment of cancer,  
 PT particularly prostate cancer.

PS Claim 24; Page; 34pp; English.

CC The present invention describes a humanised baculovirus (I) which  
 CC comprises a modified baculovirus genome having a nucleic acid molecule  
 CC encoding a therapeutic agent and a polypeptide which functions to target  
 CC the baculovirus to at least one cell type. Also described is a  
 CC pharmaceutical composition comprising (I). (I) has cytostatic activity,  
 CC and can be used in gene therapy. The baculovirus is useful in the  
 CC manufacture of a medicament for the treatment of cancer, particularly  
 CC prostate cancer. The present sequence represents the human disintegrin-  
 CC like and metalloprotease (repolysin type) with thrombospondin type 1  
 CC motif, 8 (ADMTS8) protein, which is specified in the exemplification of  
 CC the present invention as angiotensin. N.B. The present sequence is not  
 CC given in the specification but is referred to in Claim 24 as Genbank  
 CC accession number NM\_007037

XX Sequence 890 AA;

Query Match 75.8%; Score 675; DB 6; Length 890;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 875; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 14 LILLILLILLPLARGAPAPAGQASSELVPTPLPGSAGELALHLSAFKGFVRLAPDD 73  
 DB 14 LILLILLILLPLARGAPAPAGQASSELVPTPLPGSAGELALHLSAFKGFVRLAPDD 73  
 QY 14 LILLILLILLPLARGAPAPAGQASSELVPTPLPGSAGELALHLSAFKGFVRLAPDD 73  
 DB 14 LILLILLILLPLARGAPAPAGQASSELVPTPLPGSAGELALHLSAFKGFVRLAPDD 73  
 QY 74 SFLAEFKTERLGGSGRATGGERGLRGCFSGTVNGEPESLAAYSLCRLGSGSFLDGE 133  
 DB 74 SFLAEFKTERLGGSGRATGGERGLRGCFSGTVNGEPESLAAYSLCRLGSGSFLDGE 133  
 QY 134 FTIQGAGGSLAQPHRLQWGPAGARPLPRGEWEVETGEGQORERGDHDESEESQOE 193  
 DB 134 FTIQGAGGSLAQPHRLQWGPAGARPLPRGEWEVETGEGQORERGDHDESEESQOE 193  
 QY 134 FTIQGAGGSLAQPHRLQWGPAGARPLPRGEWEVETGEGQORERGDHDESEESQOE 193  
 DB 134 FTIQGAGGSLAQPHRLQWGPAGARPLPRGEWEVETGEGQORERGDHDESEESQOE 193  
 QY 194 EBAEGASEPPPLGATSRTRFVSEARFVETLLVADASAAAYGADLONHILLTMSVAA 253  
 DB 194 EBAEGASEPPPLGATSRTRFVSEARFVETLLVADASAAAYGADLONHILLTMSVAA 253  
 QY 254 IYKHPSTIKSINLMVYKVLVEDEKMGPEVSDNGGLTLNFCNQMRRTNOPSDRPHEND 313  
 DB 254 IYKHPSTIKSINLMVYKVLVEDEKMGPEVSDNGGLTLNFCNQMRRTNOPSDRPHEND 313

QY 314 TAILTRONFCGCGEGCDTLGVADIGTICDPNKSCSVIEDEGLQAANTLAHELGHVLSMP 373  
 DB 314 TAILTRONFCGCGEGCDTLGVADIGTICDPNKSCSVIEDEGLQAANTLAHELGHVLSMP 373  
 QY 374 HDSPKCTRLFGPMKHHVAPLFLVHNLQTLFMSPCSAMLTTELLOGHGDCLLDAPGA 433  
 DB 374 HDSPKCTRLFGPMKHHVAPLFLVHNLQTLFMSPCSAMLTTELLOGHGDCLLDAPGA 433  
 QY 434 LPLPTGLPGRMALVQLDQOCROI FGPDRHCNPNTSAQVCAQIMCHTDGAEPLCHTKNGS 493  
 DB 434 LPLPTGLPGRMALVQLDQOCROI FGPDRHCNPNTSAQVCAQIMCHTDGAEPLCHTKNGS 493  
 QY 494 LPMADGTPCGPGLHLCSEGSCLPEEVEVERPKPVADGMAFWGPGWGCSTCGGVQFSHRE 553  
 DB 494 LPMADGTPCGPGLHLCSEGSCLPEEVEVERPKPVADGMAFWGPGWGCSTCGGVQFSHRE 553  
 QY 554 CKDEPONGGRYCLGRARYOSCHTEBCPPDGKSFRRQOCCKYNAVYITMDGNLLQWVP 613  
 DB 554 CKDEPONGGRYCLGRARYOSCHTEBCPPDGKSFRRQOCCKYNAVYITMDGNLLQWVP 613  
 QY 614 KYAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLCPETLALCVRGQCVKAGCDHVYDS 673  
 DB 614 KYAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLCPETLALCVRGQCVKAGCDHVYDS 673  
 QY 674 PRKLDKCGVCGGKNSCKRYSGSLPTNYGYNDIVTIPAGATNDVKORSHPGVQNDGY 733  
 DB 674 PRKLDKCGVCGGKNSCKRYSGSLPTNYGYNDIVTIPAGATNDVKORSHPGVQNDGY 733  
 QY 734 LALKTADQYLLNGNLAI SAIEDIIVKGTILKYSIATLERLOSFRPLPEPLTVQLLT 793  
 DB 734 LALKTADQYLLNGNLAI SAIEDIIVKGTILKYSIATLERLOSFRPLPEPLTVQLLT 793  
 QY 794 VPGEVPPKXKTYTFVPNDVDFSMOSSKERATNTIIQPLHQAOWVLGWSBCSSTCGAG 853  
 DB 794 VPGEVPPKXKTYTFVPNDVDFSMOSSKERATNTIIQPLHQAOWVLGWSBCSSTCGAG 853  
 QY 854 QRTVECRDPSSGASATCNKALKEPADKPCESQCLPL 890  
 DB 854 QRTVECRDPSSGASATCNKALKEPADKPCESQCLPL 890

RESULT 4  
 AAB74946  
 ID AAB74946 standard; protein; 889 AA.

02-JUL-2001 (first entry)

Human ADAM type metal protease MDTs3 protein SEQ ID NO:20.

Human; a disintegrin and metalloprotease type metal protease; MDTs1;  
 MDTs2; MDTs3; ADAM type metal protease; cytosolic; antiarthritic;  
 cancer; arthritis; attheros deformans.

Homo sapiens.

JP2001008687-A.

16-JAN-2001.

25-JUN-1999; 99UP-00180973.

25-JUN-1999; 99UP-00180973.

(YAMA) YAMANOUCHI PHARM CO LTD.

WPI; 2001-285362/30.

N-PSDB; AAF82166.

New metal protease and metal protease gene, for use as a drug for  
 cancers, arthritis and attheros deformans.

XX Claim 1; Page 25-27; 31pp; Japanese.  
 PS The present sequence represents a disintegrin and metalloprotease (ADAM)  
 CC type metal protease designated MDTs3, isolated from human. MDTs proteins  
 CC have cytosolic and antiarthritic activities. They can be used as a drug  
 CC for cancers, arthritis and attheros deformans  
 XX  
 SQ Sequence 889 AA;

Query Match 72.9%; Score 649; DB 4; Length 889;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 849; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 14 LILLLLLLPLARAPAPAAAGQASSELVTRLPDGSAGELALHSAGKGFVRLAPDD 73  
 DB 13 LILLLLLLPLARAPAPAAAGQASSELVTRLPDGSAGELALHSAGKGFVRLAPDD 72  
 QY 74 SFLAEFKIERLGSGRATGERGLRGCFSGTVNGEPESLAAYSLCRGLSGSFLDGE 133  
 DB 73 SFLAEFKIERLGSGRATGERGLRGCFSGTVNGEPESLAAYSLCRGLSGSFLDGE 132  
 QY 134 FTIQOGAGGSLAOPHRLQWGPAGARLPFGPEHVEVTGCGOERGDHODESESOE 193  
 DB 133 FTIQOGAGGSLAOPHRLQWGPAGARLPFGPEHVEVTGCGOERGDHODESESOE 192  
 QY 194 EEAEGASPPPLGATSRKTFVSEARFETLLVADASMAAFYGADLQNHILTVMSVAR 253  
 DB 193 EEAEGASPPPLGATSRKTFVSEARFETLLVADASMAAFYGADLQNHILTVMSVAR 252  
 QY 254 IYKHSIKNSINLWYKYLVEDEKWPVSDNGELTRNCNMQRRNQSDDHPEHYD 313  
 DB 253 IYKHSIKNSINLWYKYLVEDEKWPVSDNGELTRNCNMQRRNQSDDHPEHYD 312  
 QY 314 TAILTRONFCGCGEGCDTLGVADIGTICDPNKSCSVIEDEGLQAANTLAHELGHVLSMP 373  
 DB 313 TAILTRONFCGCGEGCDTLGVADIGTICDPNKSCSVIEDEGLQAANTLAHELGHVLSMP 372  
 QY 374 HDSPKCTRLFGPMKHHVAPLFLVHNLQTLFMSPCSAMLTTELLOGHGDCLLDAPGA 433  
 DB 373 HDSPKCTRLFGPMKHHVAPLFLVHNLQTLFMSPCSAMLTTELLOGHGDCLLDAPGA 432  
 QY 434 LPLPTGLPGRMALVQLDQOCROI FGPDRHCNPNTSAQVCAQIMCHTDGAEPLCHTKNGS 493  
 DB 433 LPLPTGLPGRMALVQLDQOCROI FGPDRHCNPNTSAQVCAQIMCHTDGAEPLCHTKNGS 492  
 QY 494 LPMADGTPCGPGLHLCSEGSCLPEEVEVERPKPVADGMAFWGPGWGCSTCGGVQFSHRE 553  
 DB 493 LPMADGTPCGPGLHLCSEGSCLPEEVEVERPKPVADGMAFWGPGWGCSTCGGVQFSHRE 552  
 QY 554 CKDEPONGGRYCLGRARYOSCHTEBCPPDGKSFRRQOCCKYNAVYITMDGNLLQWVP 613  
 DB 553 CKDEPONGGRYCLGRARYOSCHTEBCPPDGKSFRRQOCCKYNAVYITMDGNLLQWVP 612  
 QY 614 KYAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLCPETLALCVRGQCVKAGCDHVYDS 673  
 DB 613 KYAGVSPDRCKLFCRARGSEFKVFEAKVIDGTLCPETLALCVRGQCVKAGCDHVYDS 672  
 QY 674 PRKLDKCGVCGGKNSCKRYSGSLPTNYGYNDIVTIPAGATNDVKORSHPGVQNDGY 733  
 DB 673 PRKLDKCGVCGGKNSCKRYSGSLPTNYGYNDIVTIPAGATNDVKORSHPGVQNDGY 732  
 QY 734 LALKTADQYLLNGNLAI SAIEDIIVKGTILKYSIATLERLOSFRPLPEPLTVQLLT 793  
 DB 733 LALKTADQYLLNGNLAI SAIEDIIVKGTILKYSIATLERLOSFRPLPEPLTVQLLT 792  
 QY 794 VPGEVPPKXKTYTFVPNDVDFSMOSSKERATNTIIQPLHQAOWVLGWSBCSSTCGAG 853  
 DB 793 VPGEVPPKXKTYTFVPNDVDFSMOSSKERATNTIIQPLHQAOWVLGWSBCSSTCGAG 852  
 QY 854 QRTVECRDPSS 864  
 DB 853 QRTVECRDPSS 863

```

RESULT 5
AAB21251
ID AAB21251 standard; protein; 680 AA.
AC
XX AAB21251;
AC
XX
XX 23-FEB-2001 (first entry)
XX
XX Human metalloproteinase ADAMTS-2.
DE
XX
XX Human; ADAMTS2; metalloproteinase; ADAM;
XX a disintegrin and metalloproteinase domain; thrombospondin domain;
XX vaccine; nootropic; neuroprotective; antiparkinsonian; cerebroprotective;
XX cytoskeletal; antiarthritic; immunosuppressive; Alzheimer's disease;
XX Parkinson's disease; stroke; cancer; arthritis; autoimmune disease;
XX brain tumour; brain injury.
XX
XX Homo sapiens.
OS
XX
XX WO200053774-A2.
XX
XX 14-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US006237.
XX
XX 08-MAR-1999; 99US-00264585.
XX
XX (NEUR-) NEUROCRINE BIOSCIENCES INC.
XX
XX Keiner GS, Clark M, Maki RA;
XX
XX WPI; 2000-594326/56.
XX
XX N-PSDB; AAA95821.
XX
XX Polynucleotide encoding novel members of a disintegrin, metalloproteinase
XX PT and thrombospondin domain protein family used to prevent and treat
XX Alzheimer's disease, cancer and autoimmune diseases.
XX
XX Claim 12; Fig 2; 129pp; English.
XX
XX The present sequence is human metalloproteinase ADAMTS-2. The ADAMTS
XX family of proteins is closely related to the ADAM (A Disintegrin and
XX Metalloproteinase Domain) family. Members of the ADAMTS family contain a
XX thrombospondin domain in addition to the disintegrin and
XX metalloproteinase domains found in the ADAMs. ADAMTS polypeptides are
XX useful for the manufacture of medicaments for treating conditions
XX associated with neuroinflammation and/or neurodegeneration, such as
XX Alzheimer's disease, Parkinson's disease and stroke. They are also useful
XX for treating conditions associated with cell proliferation, cell
XX migration, inflammation and/or angiogenesis, such as cancer, arthritis
XX and autoimmune diseases. They can be used to treat patients afflicted
XX with an invasive tumour, a brain tumour or brain injury.
XX
XX Sequence 680 AA;
SQ
Query Match 40.7%; Score 362; DB 3; Length 680;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 462; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 211 RTKRFVSEARFVETILVADASMAAFYGADLQNHILTMVSAARIYKPSIKNSINIMTVK 270
Db 1 RTKRFVSEARFVETILVADASMAAFYGADLQNHILTMVSAARIYKPSIKNSINIMTVK 60
QY 271 VLIYDEKMGPEVDSNGGLTLRFNCNMQRRFNQSDHPEHYDAIILTRQNFQCGQGLC 330
Db 61 VLIYDEKMGPEVDSNGGLTLRFNCNMQRRFNQSDHPEHYDAIILTRQNFQCGQGLC 120
QY 331 DTLGVADIGTICDPNKSQSVIEDEGLQAAHTLAHELGHVLSMPHDSKPTCRLFGPMGKH 390
Db 121 DTLGVADIGTICDPNKSQSVIEDEGLQAAHTLAHELGHVLSMPHDSKPTCRLFGPMGKH 180

```

```

QY 391 HMAAPLFVHNLQTLPMSPCSAMVYTELLDGHGDDLLDAPGALPLPTGLFGRMALYQLD 450
Db 181 HMAAPLFVHNLQTLPMSPCSAMVYTELLDGHGDDLLDAPGALPLPTGLFGRMALYQLD 240
QY 451 QOCROIFGDPFRHCPNPSAOPVCAQLWCHTGAEPICHTKXGSLPMAADGTPCGGHLTCE 510
Db 241 QOCROIFGDPFRHCPNPSAOPVCAQLWCHTGAEPICHTKXGSLPMAADGTPCGGHLTCE 300
QY 511 GSCLPSEEVERPKXYVDGMAFPMGPWGCSTRTGGGQVQFSHRECKDEBPONGRYCLGRR 570
Db 301 GSCLPSEEVERPKXYVDGMAFPMGPWGCSTRTGGGQVQFSHRECKDEBPONGRYCLGRR 360
QY 571 AKYOSCHTEECPPGKSPREQCEKXNAVNTDMDGNLQWPKYAGVSPDRCKLFCRA 630
Db 361 AKYOSCHTEECPPGKSPREQCEKXNAVNTDMDGNLQWPKYAGVSPDRCKLFCRA 420
QY 631 RGRSEFKYFEAKVIDGTLGPEETLAICVRGQCVAGCDHVDS 673
Db 421 RGRSEFKYFEAKVIDGTLGPEETLAICVRGQCVAGCDHVDS 463
RESULT 6
AAE09710
ID AAE09710 standard; protein; 364 AA.
XX
XX AAE09710;
XX
XX 22-NOV-2001 (first entry)
XX
XX Human gene 6 encoding novel protein HCE4D69, SEQ ID NO:57.
XX
XX Human; cytoskeletal; gene therapy; inflammatory disorder; neural disorder;
XX Parkinson's disease; Alzheimer's disease; immune system disorder; AIDS;
XX autoimmune disease; rheumatoid arthritis; muscular disorder; ischaemia;
XX reproductive disorder; Crohn's disease; pulmonary disorder; cancer;
XX myocardial infarction; glomerulonephritis; nephrotic syndrome; tumour;
XX haematopoietic disorder; rhinitis; asthma; diabetes; atherosclerosis;
XX anti-sense therapy; endocrine disorder; leukaemia.
XX
XX Homo sapiens.
OS
XX
XX WO200155202-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001325.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX
XX 07-JUL-2000; 2000US-0216880P.
XX
XX 11-JUL-2000; 2000US-0217487P.
XX
XX 11-JUL-2000; 2000US-0217496P.
XX
XX 14-JUL-2000; 2000US-0218299P.
XX
XX 26-JUL-2000; 2000US-0220963P.
XX
XX 26-JUL-2000; 2000US-0220964P.
XX
XX 14-AUG-2000; 2000US-0224518P.
XX
XX 14-AUG-2000; 2000US-0224519P.
XX
XX 14-AUG-2000; 2000US-0225213P.
XX
XX 14-AUG-2000; 2000US-0225214P.
XX
XX 14-AUG-2000; 2000US-0225266P.
XX
XX 14-AUG-2000; 2000US-0225267P.
XX
XX 14-AUG-2000; 2000US-0225268P.
XX
XX 14-AUG-2000; 2000US-0225270P.

```

[illegible]

08-NOV-2000; 2000US-0246532P.  
08-NOV-2000; 2000US-0246603P.  
08-NOV-2000; 2000US-0246610P.  
08-NOV-2000; 2000US-0246611P.  
08-NOV-2000; 2000US-0246613P.  
17-NOV-2000; 2000US-0249207P.  
17-NOV-2000; 2000US-0249208P.  
17-NOV-2000; 2000US-0249209P.  
17-NOV-2000; 2000US-0249210P.  
17-NOV-2000; 2000US-0249211P.  
17-NOV-2000; 2000US-0249212P.  
17-NOV-2000; 2000US-0249213P.  
17-NOV-2000; 2000US-0249214P.  
17-NOV-2000; 2000US-0249215P.  
17-NOV-2000; 2000US-0249216P.  
17-NOV-2000; 2000US-0249217P.  
17-NOV-2000; 2000US-0249218P.  
17-NOV-2000; 2000US-0249244P.  
17-NOV-2000; 2000US-0249245P.  
17-NOV-2000; 2000US-0249264P.  
17-NOV-2000; 2000US-0249265P.  
17-NOV-2000; 2000US-0249297P.  
17-NOV-2000; 2000US-0249299P.  
17-NOV-2000; 2000US-0249300P.  
01-DEC-2000; 2000US-0250160P.  
01-DEC-2000; 2000US-0250391P.  
05-DEC-2000; 2000US-0251030P.  
05-DEC-2000; 2000US-0251988P.  
05-DEC-2000; 2000US-0256719P.  
06-DEC-2000; 2000US-0251479P.  
08-DEC-2000; 2000US-0251856P.  
08-DEC-2000; 2000US-0251868P.  
08-DEC-2000; 2000US-0251869P.  
08-DEC-2000; 2000US-0251899P.  
08-DEC-2000; 2000US-0251990P.  
11-DEC-2000; 2000US-0254097P.  
05-JAN-2001; 2001US-0255678P.  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI: 2001-451925/48.  
N-PSDB; AAD16770.  
Isolated polypeptide for treating, preventing and/or prognosing medical disorders and also for testing and detection e.g. diagnosis and screening for agonists.  
Claim 11; SEQ ID NO 57; 469pP; English.  
AAD16750-AAD16775 represent cDNAs corresponding to novel human protein genes, and AAB09690-AAB09715 represent the proteins they encode. AAD16777-  
AAD16780 represent novel human genomic DNA fragments. The novel proteins  
and their DNAs are useful for diagnosing, treating, preventing and/or  
prognosing inflammatory disorders (bursitis or tendinitis); neural  
disorders (e.g. Parkinson's disease, Alzheimer's disease); immune system  
disorders; AIDS; autoimmune diseases (e.g., rheumatoid arthritis);  
muscular disorders; reproductive disorders; gastrointestinal disorders  
(malabsorption syndrome, Crohn's disease); pulmonary disorders;  
cardiovascular disorders (myocardial infarction, ischaemia, arrhythmias);  
renal disorders (glomerulonephritis, nephrotic syndrome); cancerous  
disease and conditions (breast cancer); hyperproliferative disorders  
(leukemia, hyperplasia); tumours; foetal and developmental abnormalities  
; haematopoietic disorders; respiratory disorders (rhinitis, asthma);  
angiogenic disorders; diabetes; atherosclerosis; endocrine disorders;  
pregnancy-related disorders and infections. The novel protein DNA is  
useful in gene therapy and anti-sense therapy. The proteins can also be  
used to aid wound healing and epithelial cell proliferation, to prevent  
skin aging due to sunburn, to maintain organs before transplantation, for  
supporting cell culture of primary tissues, to regenerate tissues, to  
identify their cognate ligands or binding partners, and in chemotaxis,  
and can be used as a food additive or preservative to modify storage

CC properties. The present sequence represents a novel human protein of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format directly  
CC from WIPO at [http://wipo.int/publ/published\\_pct\\_sequences](http://wipo.int/publ/published_pct_sequences)

Query Match 29.6%; Score 263; DB 4; Length 364;  
Best Local Similarity 100.0%; Pred. No. 4.4e-252;  
Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 530 WAPWGWGSCSRCTGGGCVGFSSRECKDPFPONGKRYCLGRRAKYOSCHTEECPPPGKSPR 589  
DB 4 WAPWGWGSCSRCTGGGCVGFSSRECKDPFPONGKRYCLGRRAKYOSCHTEECPPPGKSPR 63  
QY 590 EOCCEKXNANYNYDMGNLQWPKYAGVSPDRCLFCFRAGRESEPKYFEAKVIDGLIC 649  
DB 64 EOCCEKXNANYNYDMGNLQWPKYAGVSPDRCLFCFRAGRESEPKYFEAKVIDGLIC 123  
QY 650 GPETLAIQVGGQCVKAGCDHYVSPRLKDKCYCGKSGSKRVSGLTPITYGYNDIVT 709  
DB 124 GPETLAIQVGGQCVKAGCDHYVSPRLKDKCYCGKSGSKRVSGLTPITYGYNDIVT 183  
QY 710 IPAGATNIDVKQSHSGVQNDGNYLAKTADGQYLNGMLAISAEODILVKGTLIKTSG 769  
DB 184 IPAGATNIDVKQSHSGVQNDGNYLAKTADGQYLNGMLAISAEODILVKGTLIKTSG 243  
QY 770 SIATLERLOSFRPLPEPLTVQLL 792  
DB 244 SIATLERLOSFRPLPEPLTVQLL 266

RESULT 7  
ADCC2088  
ID ADCC2088 standard; protein; 364 AA.

XX ADCC2088;

DT 18-DEC-2003 (first entry)

DE Human protein from secreted protein gene #2.

XX Secreted protein; cytosolic; antibacterial; virucide; neuroprotective;  
KM gynascological; gastrointestinal; Gen; cardiac; cardiovascular; Gen;  
KM nephrotropic; antiinflammatory; muscular; Gen; respiratory; Gen;  
KM immunosuppressive; cerebroprotective; vasotropic; nocotropic;  
KM antiallergic; cancer; bacterial infection; viral infection;  
KM neural disorder; immune system disorder; blood disorder;  
KM muscular disorder; reproductive disorder; gastrointestinal disorder;  
KM pulmonary disorder; cardiovascular disorder; renal disorder;  
KM inflammatory disorder; proliferative disorder; Human.

XX Homo sapiens.

XX US2003082681-A1.

XX 01-MAY-2003.

PF 07-MAR-2002; 2002US-00091391.

XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 14-AUG-2000; 2000US-0224096P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225269P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225478P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232387P.  
PR 14-SEP-2000; 2000US-0232388P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 14-SEP-2000; 2000US-0233066P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235836P.  
PR 29-SEP-2000; 2000US-0236357P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 29-SEP-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.

PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249246P.  
 PR 17-NOV-2000; 2000US-0249247P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250160P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 05-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251865P.  
 PR 08-DEC-2000; 2000US-0251868P.  
 PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 PR 17-JAN-2001; 2001US-00764903.  
 (HUMA-) HUMAN GENOME SCI INC.  
 PA Rosen CA, Ruben SM, Barash SC;  
 PI WPI; 2003-786903/74.  
 DR N-PSDB; ADOC22062.  
 XX  
 PT New nucleic acid molecules and polypeptides for diagnosing, preventing or  
 PT treating disorders associated with aberrant expression of the  
 PT polypeptide, e.g. neural or cardiovascular disorders, and in chromosome  
 PT identification.  
 XX  
 PS Claim 11; SEQ ID NO 57; 242bp; English.

CC susceptibility to a pathological condition (comprising determining the  
 CC presence or amount of expression of the protein in a biological sample  
 CC and diagnosing a condition based on the presence or amount of expression  
 CC of the protein), preventing, treating or ameliorating a medical condition  
 CC by administering the nucleic acid or protein to a mammalian subject,  
 CC identifying a binding partner to the protein, the gene corresponding to  
 CC the cDNA sequence, and identifying an activity in a biological assay  
 CC (comprising expressing the nucleic acid in a cell, isolating the  
 CC supernatant, detecting an activity in a biological assay and identifying  
 CC the protein in the supernatant having the activity). The nucleic acids  
 CC and proteins display the following activities: Cytostatic, antibacterial,  
 CC Virucide, Neuroprotective, Gynaecological, Gastrointestinal, Gen,  
 CC Cardiac, Cardiovascular-Gen, Nephrotoxic, Antiinflammatory, Muscular-  
 Query Match 29.6%; Score 263; DB 7; Length 364;  
 Best Local Similarity 100.0%; Pred. No. 4,4e-252;  
 Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 530 WAPMGPMGECSTRTGGVGFHSRECKDEPONGRVCIGRAKIQSCITECPDQKSF 585  
 DB 4 WAPMGPMGECSTRTGGVGFHSRECKDEPONGRVCIGRAKIQSCITECPDQKSF 63  
 QY 590 EOCCEKNAANYTDMGDLTOWPKYAGVSPDRCKLCRARGSEEFVFAKVIDGLIC 649  
 DB 64 EOCCEKNAANYTDMGDLTOWPKYAGVSPDRCKLCRARGSEEFVFAKVIDGLIC 123  
 QY 650 GPETLAICVRGQCVAGCDHVDSERKLDKCGVCGKNSCRKVSGLTPTNYGNDIVT 709  
 DB 124 GPETLAICVRGQCVAGCDHVDSERKLDKCGVCGKNSCRKVSGLTPTNYGNDIVT 183  
 QY 710 IPAGANTIDVOKRSPGONGNYALKTADGOVILNGNLAISAIEDIILVGTILKXSG 769  
 DB 184 IPAGANTIDVOKRSPGONGNYALKTADGOVILNGNLAISAIEDIILVGTILKXSG 243  
 QY 770 SIATLERLOSFRPLPEPLTVQL 792  
 DB 244 SIATLERLOSFRPLPEPLTVQL 266  
 RESULT 8  
 AAB72285 ID AAB72285 standard; protein; 245 AA.  
 XX  
 AC AAB72285;  
 XX  
 DT 14-MAY-2001 (first entry)  
 XX  
 DE Human ADAMTS-8 amino acid sequence.  
 XX  
 KW ADAMTS-N, disintegrin; metalloprotease; thrombospondin type I motif;  
 KW tumour cachexia; inflammation; dermatosparaxis; EDS-VIC; angiogenesis;  
 KW Ehlers-Danlos syndrome type VIIC; articular cartilage erosion; human;  
 KW metacasis; embryogenesis; egg implantation; ADAMTS-8.  
 XX  
 OS Homo sapiens.  
 XX  
 PN NC0200111074-A2.  
 XX  
 PD 15-FEB-2001.  
 XX  
 PF 03-AUG-2000; 2000NC-US021223.  
 XX  
 PR 06-AUG-1999; 99US-00369364.  
 XX  
 PA (CLEV-) CLEVELAND CLINIC FOUND.  
 PA (APTE/) APTE S S.  
 PA (HURS/) HURSKATINEN T L.  
 PA (HIRO/) HIROHARA S.  
 XX  
 PI Apte SS, Hurskainen TL, Hirohata S;  
 XX WPI; 2001-159978/16.  
 DR N-PSDB; AAF63442.

XX Murine and human 'A disintegrin-like And Metalloprotease domain with  
PT Thrombospondin type I motifs' proteins and the nucleic acids encoding  
PT them, useful for treating e.g. tumors, inflammation and arthritis.  
XX  
XX  
PS Claim 1; Fig 6; 181pp; English.  
XX  
XX This invention relates to murine and human ADAMTS-N (A disintegrin-like  
CC and metalloprotease domain with thrombospondin type I motifs) proteins,  
CC designated ADAMTS-5, 6, 7, 8, 9, 10 and R1. Also included in the  
CC invention are cDNA sequences encoding the proteins, and antibodies  
CC specific for the proteins. The nucleic acid sequences and proteins may be  
CC used in the prevention, diagnosis and treatment of diseases associated  
CC with inappropriate ADAMTS-N expression. Disorders that may be treated  
CC using the nucleic acids, proteins and antibodies include, for example  
CC tumour cachexia, inflammation, dermatosparaxis in cattle or Ehlers-Danlos  
CC syndrome type VIIC (EDS-VIIC) in humans, erosion of articular cartilage  
CC in arthritic (both inflammatory and non-inflammatory) disease,  
CC angiogenesis, tumour growth and metastases, and they may also be used for  
CC controlling embryogenesis and implantation of fertilised eggs. The  
CC present sequence represents human ADAMTS-8  
XX  
SQ Sequence 245 AA:  
Query Match 24.4%; Score 217; DB 4; Length 245;  
Best Local Similarity 100.0%; Pred. No. 1.6e-206;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 196 AEGASEPPPLGATSRTRKFEVSEARFVETLLVADASMAAFYAGDLQNHITLMSVARIY 255  
2 AEGASEPPPLGATSRTRKFEVSEARFVETLLVADASMAAFYAGDLQNHITLMSVARIY 61  
QY 256 KHPISKISINLMVYKVILVDEKMGKPEVSDNGGTLTLPNCRNRPNQSDRHPHYDTA 315  
DB 62 KHPISKISINLMVYKVILVDEKMGKPEVSDNGGTLTLPNCRNRPNQSDRHPHYDTA 121  
QY 316 ILLTRONFCGQEGCDTLGVADIGTICDPNKSQVIEDEGLQAHTLAHGLVLSMPHD 375  
DB 122 ILLTRONFCGQEGCDTLGVADIGTICDPNKSQVIEDEGLQAHTLAHGLVLSMPHD 181  
QY 376 DSKPCTRLFGPMGKHVMAPLFVHLNQLTLPWSPCSAM 412  
DB 182 DSKPCTRLFGPMGKHVMAPLFVHLNQLTLPWSPCSAM 218  
RESULT 9  
AAB72284  
ID AAB72284 standard; protein; 905 AA.  
XX  
XX AAB72284;  
AC  
XX  
XX 14-MAY-2001 (first entry)  
DT  
XX  
XX Murine ADAMTS-8 amino acid sequence.  
DE  
XX  
XX ADAMTS-N; disintegrin; metalloprotease; thrombospondin type I motif;  
KW tumour cachexia; inflammation; dermatosparaxis; EDS-VIIC; angiogenesis;  
KW Ehlers-Danlos syndrome type VIIC; articular cartilage erosion; mouse;  
KW metastasis; embryogenesis; egg implantation; ADAMTS-8.  
KM  
XX  
XX Mus musculus.  
OS  
XX  
XX WO20011074-A2.  
PN  
XX  
XX 15-FEB-2001.  
PD  
XX  
XX 03-AUG-2000; 2000WO-US021223.  
PF  
XX  
XX 06-AUG-1999; 99US-00369364.  
PR  
XX  
XX (CLEV-) CLEVELAND CLINIC FOUND.  
PA (APTE-) APTE S S.  
PA (HURS/) HURSKAINEN T L.

PA (HIRO/) HIROHATA S.  
XX  
XX Apte SS, Hurskainen TL, Hirohata S;  
PI  
XX  
XX WPI; 2001-159978/16.  
DR  
XX  
XX N-PSDB; AAF63441.  
DR  
XX  
XX  
PT Murine and human 'A Disintegrin-like And Metalloprotease domain with  
PT Thrombospondin type I motifs' proteins and the nucleic acids encoding  
PT them, useful for treating e.g. tumors, inflammation and arthritis.  
XX  
XX  
PS Claim 1; Fig 5; 181pp; English.  
XX  
XX This invention relates to murine and human ADAMTS-N (A disintegrin-like  
CC and metalloprotease domain with thrombospondin type I motifs) proteins,  
CC designated ADAMTS-5, 6, 7, 8, 9, 10 and R1. Also included in the  
CC invention are cDNA sequences encoding the proteins, and antibodies  
CC specific for the proteins. The nucleic acid sequences and proteins may be  
CC used in the prevention, diagnosis and treatment of diseases associated  
CC with inappropriate ADAMTS-N expression. Disorders that may be treated  
CC using the nucleic acids, proteins and antibodies include, for example  
CC tumour cachexia, inflammation, dermatosparaxis in cattle or Ehlers-Danlos  
CC syndrome type VIIC (EDS-VIIC) in humans, erosion of articular cartilage  
CC in arthritic (both inflammatory and non-inflammatory) disease,  
CC angiogenesis, tumour growth and metastases, and they may also be used for  
CC controlling embryogenesis and implantation of fertilised eggs. The  
CC present sequence represents murine ADAMTS-8  
XX  
SQ Sequence 905 AA:  
Query Match 3.9%; Score 35; DB 4; Length 905;  
Best Local Similarity 100.0%; Pred. No. 3.8e-25;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 617 GVSPPDRCKLFCRARGSEPFVFAKVIDGTLGCP 651  
632 GVSPPDRCKLFCRARGSEPFVFAKVIDGTLGCP 666  
RESULT 10  
AA04145  
ID AA04145 standard; protein; 481 AA.  
XX  
XX AA04145;  
AC  
XX  
XX 15-JUN-1999 (first entry)  
DT  
XX  
XX Rat Tango-76 protein.  
DE  
XX  
XX Rat; Tango-71; Tango-73; Tango-74; Tango-76; Tango-83; diagnosis;  
KW detection.  
KW  
XX  
XX Rattus sp.  
OS  
XX  
XX WO9907850-A1.  
PN  
XX  
XX 18-FEB-1999.  
PD  
XX  
XX 06-AUG-1998; 98WO-US016502.  
PF  
XX  
XX 06-AUG-1997; 97US-0054966P.  
PR  
XX  
XX 05-SEP-1997; 97US-0058108P.  
XX  
XX (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.  
PA  
XX  
XX Holtzman DA, Goodearl ADJ;  
PI  
XX  
XX WPI; 1999-167426/14.  
DR  
XX  
XX N-PSDB; AAX19958.  
DR  
XX  
XX New TANGO polypeptides and nucleic acids encoding them - useful as  
PT diagnostic agents and for treating disorders caused by aberrant  
PT expression of TANGO.  
PT





CC Cancer metaeases  
XX  
SQ Sequence 2150 AA;

Query Match 1.5%; Score 13; DB 3; Length 2150;  
Best Local Similarity 100.0%; Pred. No. 0.0061;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGAVQ 548  
|||  
Db 597 WGECSRTCGGAVQ 609

RESULT 13  
AAB90617  
ID AAB90617 standard; protein; 2165 AA.  
XX  
AC AAB90617;  
XX  
DT 01-JUN-2001 (first entry)  
XX  
DE Human secreted protein, SEQ ID NO: 155.  
XX  
KW Human; secreted protein; immunomodulatory; antisclerotic; dermatological;  
KW antiinflammatory; anti-HIV; cytostatic; cardiac; vascular;  
KW anti-angiogenic; ophthalmological; neuroprotectant; nootropic;  
KW anticonvulsant; antialzheimers; antiparkinsonian; antitubercial;  
KW vulnerary; vaccine; gene therapy; cancer; protein coordinate data;  
infection.  
XX  
OS Homo sapiens.  
XX  
PN WC200121658-A1.  
XX  
PD 29-MAR-2001.  
XX  
PF 22-SEP-2000; 2000MO-US026013.  
XX  
PR 24-SEP-1999; 99US-0155709P.  
XX  
PA (HUMA-) HUMAN GENOME SCT INC.  
XX  
PI Ni J, Baker KP, Birse CE, Ebner R, Fisceella M, Komatsoulis GA;  
PI Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;  
PI Young PE, Wei P, Florence KA;  
XX  
DR WPI; 2001-235311/24.

Nucleic acids encoding 32 human secreted polypeptides, useful for preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease and diabetic retinopathy.

Disclousre; Page 830-836; 890pp; English.

The present sequence is provided in a specification relating to nucleic acid molecules encoding 32 novel human secreted polypeptides. The nucleic acid molecules and polypeptides may be used in the prevention, diagnosis and treatment of diseases such as immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus and human immuno-deficiency virus (HIV) infections), hyperproliferative disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases (e.g. Schmitz syndrome, Chaga's cardiomyopathy and coronary arteriosclerosis), angio-genic disorders (e.g. corneal graft neovascularisation and diabetic retinopathy), neurological disorders (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease), infectious diseases and/or for promoting wound healing, regeneration and/or chemotaxis. The nucleic acid molecules may be used to produce the secreted polypeptides. They may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples. The polypeptides may be used as antigens in the production of antibodies and in assays to identify modulators of their expression and activity

Sequence 2165 AA;

Query Match 1.5%; Score 13; DB 4; Length 2165;  
Best Local Similarity 100.0%; Pred. No. 0.0061;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGAVQ 548  
|||  
Db 612 WGECSRTCGGAVQ 624

RESULT 14  
AAP81021  
ID AAP81021 standard; protein; 18 AA.  
XX  
AC AAP81021;  
XX  
DT 15-NOV-1990 (first entry)  
XX  
DE Sequence of modified egg white lysozyme signal peptide.  
XX  
KW Modified egg white lysozyme signal sequence; enzyme.  
XX  
OS Homo sapiens.  
XX  
PN JP63233789-A.  
XX  
PD 29-SEP-1988.  
XX  
PF 23-MAR-1987; 87JP-00069764.  
XX  
PR 23-MAR-1987; 87JP-00069764.  
XX  
PA (AJIN ) TANPAKU KAGAKU KENK.  
XX  
DR WPI; 1988-318076/45.  
XX  
PD N-PSDB; AAN81348.  
XX  
PT DNA sequence encoding signal peptide - having good protein secretion activity.  
XX  
PS Claim 3; Page 523; 11pp; Japanese.  
XX  
CC AAs of natural egg white lysozyme other than residues 1, 2 and 16-18 are  
CC all or mostly substituted with one kind of hydrophobic AA. Using the  
CC modified signal peptide, foreign proteins, esp. human lysozyme, can be  
CC secreted correctly and at a high level  
XX  
SQ Sequence 18 AA;

Query Match 1.3%; Score 12; DB 1; Length 18;  
Best Local Similarity 100.0%; Pred. No. 0.00082;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLLPLA 25  
|||  
Db 4 LLLLLLLLLPLA 15

RESULT 15  
AAB97107  
ID AAB97107 standard; peptide; 19 AA.  
XX  
AC AAB97107;  
XX  
DT 02-AUG-2001 (first entry)  
XX  
DE Chlorella signal peptide.  
XX  
KW Chlorella; signal peptide; gene expression; protein production;  
XX human growth hormone.  
XX  
OS Chlorella sp.  
XX

PN JP2000354490-A.  
XX 26-DEC-2000.  
PD  
XX 15-JUN-1999; 99JP-00168271.  
PF  
XX 15-JUN-1999; 99JP-00168271.  
PR  
XX (TOYT ) TOYOTA JIDOSHA KK.  
PA  
XX WPI; 2001-275809/29.  
DR  
XX New signal peptides useful for the preparation of human growth hormone  
PT and transformed chlorella.  
PT  
XX Disclosure; Page 12; 15pp; Japanese.  
PS  
XX The present sequence is provided in a specification relating to signal  
CC peptides for expression and secretion of a protein in chlorella. The  
CC peptides are of the formula: Met-Ala-Asn-Lys-X<sub>1</sub>-(Leu)-n-X<sub>2</sub>-Ala-Ser-Gly.  
CC X<sub>1</sub> = Ser or Leu; n = an integer of 5-15; X<sub>2</sub> = Gly-Ser-Leu or Pro-Leu-  
CC Ala. The signal peptides are useful in the preparation of human growth  
CC hormone and transformed chlorella. Signal peptides, DNA encoding the  
CC peptides, gene expression cassettes, recombinant vectors containing the  
CC cassettes, and transformants having the vectors are provided  
XX  
SQ Sequence 19 AA;  
Query Match 1.3%; Score 12; DB 4; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.00086;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 14 LLLLLLLLLPLA 25  
DB 5 LLLLLLLLLPLA 16

Search completed: May 7, 2004, 11:58:36  
Job time : 68 secs

Fri May 7 12:18:02 2004

us-09-989-687-4\_1.rapb

Page 1

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:51:11 ; Search time 54 Seconds  
(without Alignment)  
4574.711 Million cell updates/sec

Title: US-09-989-687-4  
Sequence: 1 MFPAAPAPWLPFLILLILL.....CNKALKPDAKPCESQCLPL 890

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1140673 segs, 277566755 residues

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*  
1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubppaa/FCI\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep.\*  
6: /cgn2\_6/ptodata/2/pubppaa/PCITUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4853	100.0	890	10	US-09-373-658-4
2	4853	100.0	890	11	US-09-989-687-4
3	4810	99.1	924	12	US-10-425-114-39107
4	3919.5	80.8	905	9	US-09-918-171A-9
5	2344	48.3	950	9	US-09-321-987B-4
6	2344	48.3	950	13	US-10-163-316-7
7	2328.5	48.0	950	11	US-09-373-658-2
8	2328.5	48.0	950	11	US-09-989-687-2
9	2328.5	48.0	967	13	US-10-105-929-2
10	2328.5	48.0	967	14	US-10-115-286-2
11	2328.5	48.0	968	10	US-09-373-658-125
12	2328	47.8	931	12	US-09-741-151-4
13	2320.5	47.8	967	11	US-09-989-687-126
14	2274.5	46.9	481	12	US-09-802-582-8
15	2274.5	46.9	481	13	US-10-105-929-8

16	2274.5	46.9	481	14	US-10-365-227-8	Sequence 8, Appl
17	2146.5	44.2	924	15	US-10-093-463-28	Sequence 28, Appl
18	2138.5	44.1	950	9	US-09-965-631-4	Sequence 4, Appl
19	2138.5	44.1	950	12	US-09-741-151-2	Sequence 2, Appl
20	2128.5	43.9	978	12	US-10-275-107-59	Sequence 59, Appl
21	2115.5	43.6	952	16	US-10-311-035-11	Sequence 11, Appl
22	2103.5	43.3	727	9	US-09-445-023A-1	Sequence 1, Appl
23	2103.5	43.3	727	14	US-10-097-597-1	Sequence 1, Appl
24	2103.5	43.3	727	14	US-10-097-580-1	Sequence 12, Appl
25	2093.5	43.1	727	14	US-09-445-023A-12	Sequence 12, Appl
26	2093.5	43.1	727	14	US-10-097-580-12	Sequence 12, Appl
27	2093.5	43.1	822	13	US-10-163-316-2	Sequence 2, Appl
28	1989.5	41.0	364	9	US-09-764-903-57	Sequence 57, Appl
29	1842.5	38.0	837	12	US-10-358-283-11	Sequence 11, Appl
30	1842.5	38.0	837	14	US-10-247-685-2	Sequence 2, Appl
31	1842.5	37.9	837	10	US-09-946-374-317	Sequence 317, Appl
32	1841.5	37.9	837	12	US-10-206-915-352	Sequence 352, Appl
33	1841.5	37.9	837	12	US-10-199-670-352	Sequence 352, Appl
34	1841.5	37.9	837	12	US-10-201-858-352	Sequence 352, Appl
35	1841.5	37.9	837	12	US-10-205-890-352	Sequence 352, Appl
36	1841.5	37.9	837	12	US-10-208-024-352	Sequence 352, Appl
37	1841.5	37.9	837	12	US-10-201-853-352	Sequence 352, Appl
38	1841.5	37.9	837	12	US-10-174-581-352	Sequence 352, Appl
39	1841.5	37.9	837	12	US-10-174-483-352	Sequence 352, Appl
40	1841.5	37.9	837	12	US-10-176-914-352	Sequence 352, Appl
41	1841.5	37.9	837	12	US-10-176-914-352	Sequence 352, Appl
42	1841.5	37.9	837	12	US-10-176-915-352	Sequence 352, Appl
43	1841.5	37.9	837	12	US-10-006-485A-317	Sequence 317, Appl
44	1841.5	37.9	837	12	US-10-013-907A-317	Sequence 317, Appl
45	1841.5	37.9	837	12	US-10-013-907A-317	Sequence 317, Appl

# ALIGNMENTS

RESULT 1  
US-09-373-658-4  
Sequence 4, Application US/09373658  
Publicatation No US20030092900A1  
GENERAL INFORMATION:  
APPLICANT: Iruela-Arispe, Luisa  
APPLICANT: Hastings, Gregg A.  
APPLICANT: Ruben, Steven M.  
APPLICANT: Jonak, Zdenka L.  
APPLICANT: Trulli, Stephen H.  
APPLICANT: Fromwald, James A.  
APPLICANT: Terrett, Jonathan A.  
TITLE OF INVENTION: Mech1 and Mech2 Polynucleotides and Polypeptides  
FILE REFERENCE: 1488.1070006  
CURRENT APPLICATION NUMBER: US/09/373,658  
CURRENT FILING DATE: 1999-08-13  
NUMBER OF SEQ ID NOS: 125  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 890  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-373-658-4  
Query Match 100.0%; Score 4853; DB 10; Length 890;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MFPAAPAPWLPFLILLILLPLARGAAPARAAGGASLVVPRTPSSAGEILATHSA 60  
DB 1 MFPAAPAPWLPFLILLILLPLARGAAPARAAGGASLVVPRTPSSAGEILATHSA 60  
QY 61 FGAGFLRLAPADPSFLAPFEXIERGGSGRATGGERGLRGCFSGTVNGEPSSLAIVSLC 120  
DB 61 FGAGFLRLAPADPSFLAPFEXIERGGSGRATGGERGLRGCFSGTVNGEPSSLAIVSLC 120  
QY 121 RGISGFLDGEFFTPQAGAGSLAOPHRLQWAGARPLPRGSEVENVEGEGOROR 180  
DB 121 RGISGFLDGEFFTPQAGAGSLAOPHRLQWAGARPLPRGSEVENVEGEGOROR 180

Dh	122	RGJSSGFLIDGBEFTIQOGAGGSJAOQHRIQOFMGPGAPPLPRGSEMEYETEGGROER	180
Qy	181	GHOHDSSESESEBEAEAGSEPPPLGATSTTXFVESEAPFVETLLVAASMAAFGADL	240
Dh	181	GHOHEESESESEBEAGSEPPPLGATSTTXFVESEAPFVETLLVAASMAAFGADL	240
Qy	241	QNHILLMSVARIYKHPSIKNSILMLMVKYLVEDEKMGBEVSDNGILLTRNFCNMQR	300
Dh	241	QNHILLMSVARIYKHPSIKNSILMLMVKYLVEDEKMGBEVSDNGILLTRNFCNMQR	300
Qy	301	FNOPSREHEHTDAILLTRONFCQOEGICDTLGVADILGITOPNKSCTYIEBEGIOAH	360
Dh	301	FNOPSREHEHTDAILLTRONFCQOEGICDTLGVADILGITOPNKSCTYIEBEGIOAH	360
Qy	361	TLNHEIGHYLSMPEHDSKPCTRLFPGMKHHVMAPLFVHLNQTLPMSPCSAMYLTELDDG	420
Dh	361	TLNHEIGHYLSMPEHDSKPCTRLFPGMKHHVMAPLFVHLNQTLPMSPCSAMYLTELDDG	420
Qy	421	GHEDCILLDAPGALPFTGLPGRMALYOLDQOCSRIEPPRHRHEPMSADVDCAOLMCHT	480
Dh	421	GHEDCILLDAPGALPFTGLPGRMALYOLDQOCSRIEPPRHRHEPMSADVDCAOLMCHT	480
Qy	481	DGAEPLCHTKNGSLPMADGTPCGSPHLCSESGCLPFESEVERPKFVVUDGVAWPGWGECS	540
Dh	481	DGAEPLCHTKNGSLPMADGTPCGSPHLCSESGCLPFESEVERPKFVVUDGVAWPGWGECS	540
Qy	541	RTCGGGVQPSHRECKDPEPONGGARYCLGHRATYQSCHEECPPRGKSFRRQOCEKXNAYN	600
Dh	541	RTCGGGVQPSHRECKDPEPONGGARYCLGHRATYQSCHEECPPRGKSFRRQOCEKXNAYN	600
Qy	601	YTTMDGNLLOWVEKYAGVSPRDRCKLFCARARSSEFKFEAKVLDGLCPETLAIQVRG	660
Dh	601	YTTMDGNLLOWVEKYAGVSPRDRCKLFCARARSSEFKFEAKVLDGLCPETLAIQVRG	660
Qy	661	QCYKAGCDHYVDSPKLIDKCYGCGGKXNSCRVSSSLPPTYNGNDIVTIPAGATINDVK	720
Dh	661	QCYKAGCDHYVDSPKLIDKCYGCGGKXNSCRVSSSLPPTYNGNDIVTIPAGATINDVK	720
Qy	721	QSHSGPQNDGNVLAKTADGQVTLNGNLAISAEODILVKGTLIKTSGSIALTERLOSJ	780
Dh	721	QSHSGPQNDGNVLAKTADGQVTLNGNLAISAEODILVKGTLIKTSGSIALTERLOSJ	780
Qy	781	RPIEPLTVLLTVPGSEVPPPKYKTYTFPVENDVDPSMOSSXERATINIQLPLHAQWVLG	840
Dh	781	RPIEPLTVLLTVPGSEVPPPKYKTYTFPVENDVDPSMOSSXERATINIQLPLHAQWVLG	840
Qy	841	DMSCSSTCGAGMORRTVECRDPSGQASATCNKALKEPDAPCESQJLCPJ	890
Dh	841	DMSCSSTCGAGMORRTVECRDPSGQASATCNKALKEPDAPCESQJLCPJ	890

```

RESULT 2
US-09-989-687-4
; Sequence 4, Application US/09989687
; Publication No. US20040002449A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1468.107000D
; CURRENT APPLICATION NUMBER: US/09/989,687
; CURRENT FILING DATE: 2001-11-21
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 890
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-989-687-4

Query Match      100.0%   Score 4853; DB 11; Length 890;
Best Local Similarity 100.0%   Pred. No. 0;
Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	1	MFEPAPRPMPLFLLILLILLILLIPARAPAPFAAGGQASELIVTRLLSGSAGELALHISA	60
Db	1	MFEPAPRPMPLFLLILLILLILLIPARAPAPFAAGGQASELIVTRLLSGSAGELALHISA	60
QY	61	FGKGFVRLAPDDSFLLAPFKIERLLGSGGRATGGERGLRGCFEFGTVNGEESLAASVLC	120
Db	61	FGKGFVRLAPDDSFLLAPFKIERLLGSGGRATGGERGLRGCFEFGTVNGEESLAASVLC	120
QY	121	RGLGSGFLDDEEFTIOPQAGAGSLAOPHRLOKWPAGAPARLPFGREWEVETGSGORCER	180
Db	121	RGLGSGFLDDEEFTIOPQAGAGSLAOPHRLOKWPAGAPARLPFGREWEVETGSGORCER	180
QY	181	GDHODESEESOEERAEASBPPLIATSTTKFVFSARFVETLLVADASMAARYADL	240
Db	181	GDHODESEESOEERAEASBPPLIATSTTKFVFSARFVETLLVADASMAARYADL	240
QY	241	QNHILLTMSVAAATYKHPBSIKNSINLMVVKYLIVDEKMGPEVSDNGELTLRNFENMOR	300
Db	241	QNHILLTMSVAAATYKHPBSIKNSINLMVVKYLIVDEKMGPEVSDNGELTLRNFENMOR	300
QY	301	FNQSDRREHYDPAILLTRQNFQCGEELCVTLGYADIGTICDPKKSCSVIEDSGLOAAH	360
Db	301	FNQSDRREHYDPAILLTRQNFQCGEELCVTLGYADIGTICDPKKSCSVIEDSGLOAAH	360
QY	361	TLAHELGHVLSMPHDSKPCRTLFGPMGKHVMAVLFVHNLQTLFWSPCSANYLLELLDG	420
Db	361	TLAHELGHVLSMPHDSKPCRTLFGPMGKHVMAVLFVHNLQTLFWSPCSANYLLELLDG	420
QY	421	GHGDCILAPGAAALPPLTGLFGPAAALYOLDQOCROIFEPDFRHCNTSADQVCAOLMCHT	480
Db	421	GHGDCILAPGAAALPPLTGLFGPAAALYOLDQOCROIFEPDFRHCNTSADQVCAOLMCHT	480
QY	481	DGAEPULCHTKNGSLFWADGTCGPGPHLCSBSGLPBEVEVERKPVVDGMAWGMWGBGS	540
Db	481	DGAEPULCHTKNGSLFWADGTCGPGPHLCSBSGLPBEVEVERKPVVDGMAWGMWGBGS	540
QY	541	RTCGGAGVQFSHRECDPEPONGARCYCLPRRAYQSCHTEECPDPDGSFPEOOCSEKYNAYN	600
Db	541	RTCGGAGVQFSHRECDPEPONGARCYCLPRRAYQSCHTEECPDPDGSFPEOOCSEKYNAYN	600
QY	601	YTDMDGNLLOWPKYAGVSPDRCLFCARARSFVKYFEAVIDIGTLGPEFLAIQVYG	660
Db	601	YTDMDGNLLOWPKYAGVSPDRCLFCARARSFVKYFEAVIDIGTLGPEFLAIQVYG	660
QY	661	QCVKAGCHVVDSPFKLDKCGVCGGKGNSCRKVSGLTFTNYGYNDIVTIPAGATNIDVK	720
Db	661	QCVKAGCHVVDSPFKLDKCGVCGGKGNSCRKVSGLTFTNYGYNDIVTIPAGATNIDVK	720
QY	721	QESHGQVQNDGNYLALKTADGQYLLANGNLAISAIEDILLVKYSSINTLELQSF	780
Db	721	QESHGQVQNDGNYLALKTADGQYLLANGNLAISAIEDILLVKYSSINTLELQSF	780
QY	781	RLPEPELTQVLLTVEGEVFPFKYKTYTFVNDVDFSMOSSKERATNTIIOPLIHAQWVLG	840
Db	781	RLPEPELTQVLLTVEGEVFPFKYKTYTFVNDVDFSMOSSKERATNTIIOPLIHAQWVLG	840
QY	841	DMSBGSSTGAGMORRTVCRDPSGQASATCNKALKPEDAKCESQLCPL	890
Db	841	DMSBGSSTGAGMORRTVCRDPSGQASATCNKALKPEDAKCESQLCPL	890

```

RESULT 3
US-10-425-114-39107
; Sequence 39107, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingsong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei

```

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53313) B  
CURRENT APPLICATION NUMBER: US/10/425,114  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ. ID NOS: 73128  
SEQ. ID NO 39107  
LENGTH: 924  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Clone ID: LIB4654-025-F6\_P1.pcp  
US-10-425-114-39107

Query Match 99.1%; Score 4810; DB 12; Length 924;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 884; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

1 MFPAAPRMFLPLLLLLLPLAAGAPAPPAAGGASELVPTPLPGSAGELALHLSA 60  
37 MLPAAPRMFP-LLLLLLPLAAGAPAPPAAGGASELVPTPLPGSAGELALHLSA 94  
61 FGKGFVRLAAPDSDFLAPEFKIERLGGSGRATGGERLRCGCFSGTVNGEPESLAIVSLC 120  
95 FGKGFVRLAAPDSDFLAPEFKIERLGGSGRATGGERLRCGCFSGTVNGEPESLAIVSLC 154  
121 RGLSGSFLDGEETTPQAGAGSLAOPHRLORNGPAGAPLPRGPEWETEGEGORER 180  
155 RGLSGSFLDGEETTPQAGAGSLAOPHRLORNGPAGAPLPRGPEWETEGEGORER 214  
181 GDHODESEESQEEAEAGASPPPLGATSTRKFSVSEARFVETLLVADASMAAFYADL 240  
215 GDHODESEESQEEAEAGASPPPLGATSTRKFSVSEARFVETLLVADASMAAFYADL 274  
241 QNHLLTMSVAARIYKPSIKNSINLMVVKYLVEDEKMGPEVSDNGLLTRAFCKMQR 300  
275 QNHLLTMSVAARIYKPSIKNSINLMVVKYLVEDEKMGPEVSDNGLLTRAFCKMQR 334  
301 FNQSDRPEHYDAIILTRONFCQEGELCDTLGVADIGTICDPMKCSVIEDEGLQAAH 360  
335 FNQSDRPEHYDAIILTRONFCQEGELCDTLGVADIGTICDPMKCSVIEDEGLQAAH 394  
361 TLAEHLGVLNMPHDSKPCRTLFGPMKHHVMAPLFHLNQLTPMSPCSAMYTELLDG 420  
395 TLAEHLGVLNMPHDSKPCRTLFGPMKHHVMAPLFHLNQLTPMSPCSAMYTELLDG 454  
421 GHGDCLLAPGAALPLPTGLPGRMALYQLDOCCROIFGPDFRHCENISAQVCAQLWCHT 480  
455 GHGDCLLAPGAALPLPTGLPGRMALYQLDOCCROIFGPDFRHCENISAQVCAQLWCHT 514  
481 DGAEPDLCTKNGSLPMADGTPCGPHLCSGSCLEPEEVEPRKPYVDGMAFWMGWCS 540  
515 DGAEPDLCTKNGSLPMADGTPCGPHLCSGSCLEPEEVEPRKPYVDGMAFWMGWCS 574  
541 RCGGGVQFSHRECDPEPONGRCYCLGRPAKYOSCHTEECPPDGSFRECCCKYANN 600  
575 RCGGGVQFSHRECDPEPONGRCYCLGRPAKYOSCHTEECPPDGSFRECCCKYANN 634  
601 YTDMDGNLLQWPKTAGVSPDRCKLFCRARGSEFKYFEAKVIDGTLGPEPTLAIYVRG 660  
635 YTDMDGNLLQWPKTAGVSPDRCKLFCRARGSEFKYFEAKVIDGTLGPEPTLAIYVRG 694  
661 QCVKAGCHVVDSPKLDKCGVCGKNGSCRYGSLPTNYGVNDYITIPAGATNIDYK 720  
695 QCVKAGCHVVDSPKLDKCGVCGKNGSCRYGSLPTNYGVNDYITIPAGATNIDYK 754  
721 QSHSPVQNDGNVYALKTADQGYLLNGNLAIISAIEDIIVKGTILKYSGSITATLERSF 780  
755 QSHSPVQNDGNVYALKTADQGYLLNGNLAIISAIEDIIVKGTILKYSGSITATLERSF 814  
781 RPLPELTYQLTVPESEVPPKPYTFPNDVDSMOSKERRATNTNIIQPLILHQLWLG 840  
815 RPLPELTYQLTVPESEVPPKPYTFPNDVDSMOSKERRATNTNIIQPLILHQLWLG 874

841 DMSRCSSTCGAGMORATVECRDPGASATCNKALKEPDAKPCESQUCPL 890  
875 DMSRCSSTCGAGMORATVECRDPGASATCNKALKEPDAKPCESQUCPL 924

RESULT 4  
US-09-918-171A-9  
Sequence 9, Application US/09918171A  
Patent No. US20020110894A1  
GENERAL INFORMATION:  
APPLICANT: Apte, Suneel  
APPLICANT: Hurskainen, Titta L.  
APPLICANT: Hirohata, Satoru  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/04193  
CURRENT APPLICATION NUMBER: US/09/918,171A  
CURRENT FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 09/369,364  
PRIOR FILING DATE: 1999-08-06  
NUMBER OF SEQ. ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ. ID NO 9  
LENGTH: 905  
TYPE: PRT  
ORGANISM: Mus musculus ADAMTS-8  
US-09-918-171A-9

Query Match 80.8%; Score 3919.5; DB 9; Length 905;  
Best Local Similarity 79.2%; Pred. No. 5,4e-303;  
Matches 717; Conservative 74; Mismatches 99; Indels 15; Gaps 4;

1 MFPAAPRMFLPLLLLLLPLAAGAPAPPAAGGASELVPTPLPGSAGELALHLS 59  
1 MRDPTTGWPLPLLLLLLQPPPLVCAAPGPGTGAQASELVPTPLPGSAGELALHLS 60  
60 AFGKGFVRLAAPDSDFLAPEFKIERLGGSGRATGGERLRCGCFSGTVNGEPESLAIVSL 119  
61 AFGKGFVRLAAPDSDFLAPEFKIERLGGSGRATGGERLRCGCFSGTVNGEPESLAIVSL 120  
120 CRGLSGSFLDGEETTPQAGAGSLAOPHRLORNGPAGAPLPRGPEWETEGEGORER 168  
121 VAGMSGSFLDGEETTPQAGAGSLAOPHRLORNGPAGAPLPRGPEWETEGEGORER 180  
169 EVETGRCQOEGRDHOESESQEEAEAGASPPPLGATSTRKFSVSEARFVETLL 226  
181 EVETGRCQOEGRDHOESESQEEAEAGASPPPLGATSTRKFSVSEARFVETLL 240  
227 VADASMAAFYADLQNHLLTMSVAARIYKPSIKNSINLMVVKYLVEDEKMGPEVSDN 286  
241 VADASMAAFYADLQNHLLTMSVAARIYKPSIKNSINLMVVKYLVEDEKMGPEVSDN 300  
287 GGLTLNFCMQRFPNQBDRPEHYDAIILTRONFCQEGELCDTLGVADIGTICDPMK 346  
301 GGLTLNFCMQRFPNQBDRPEHYDAIILTRONFCQEGELCDTLGVADIGTICDPMK 360  
347 SCSEVIEDEGLQAAHTLAHELGHVLSMPHDSKPCRTLFGPMKHHVMAPLFHLNQLTP 406  
361 SCSEVIEDEGLQAAHTLAHELGHVLSMPHDSKPCRTLFGPMKHHVMAPLFHLNQLTP 420  
407 SPSCAMTLLTLDGGHDCCLDAPGAALPLPTGLPGRMALYQLDOCCROIFGPDFRHC 466  
421 SPSCAMTLLTLDGGHDCCLDAPGAALPLPTGLPGRMALYQLDOCCROIFGPDFRHC 480  
467 TSAQDVCAQLMC-HTDGAEPDLCTKNGSLPMADGTPCGPHLCSGSCLEPEEVEPRK 525  
481 TSAQDVCAQLMC-HTDGAEPDLCTKNGSLPMADGTPCGPHLCSGSCLEPEEVEPRK 540  
526 VDGGMAFWMGWCSKRCYCLGRPAKYOSCHTEECPPDGSFRECCCKYANN 585  
541 VDGGMAFWMGWCSKRCYCLGRPAKYOSCHTEECPPDGSFRECCCKYANN 600  
586 KSFRECCCKYANNYTDMDGNLLQWPKTAGVSPDRCKLFCRARGSEFKYFEAKVID 645

Db 601 KSPFQCEKYNVNHDLGNLQWPKYSGVSPDRCKLFCRARGSEFVFAKXID 660  
 Qy 646 GTLCGPETLACVRGQCVKAGCHVYDSPKLDKCGVCGKNSCRKXSGSLTPNNYGIN 705  
 Db 661 GTLCGPETLACVRGQCVKAGCHVYDSPKLDKCGVCGKNSCRKXSGSLTPNNYGIN 720  
 Qy 706 DIWTPAGATNIDVKQSHPGVQNDGNVYALKTADQVLLNGNLAIASIEODIYKGTIL 765  
 Db 721 DIWTPAGATNIDVKQSHPGVQNDGNVYALKTADQVLLNGNLAIASIEODIYKGTIL 780  
 Qy 766 KYGSGSTLTERLQSFRLPEPLVYVOLLTPGVEFPKXKTYTFVNDVDFMSQSKERAT 825  
 Db 781 KYGSGSTLTERLQSFRLPEPLVYVOLLTPGVEFPKXKTYTFVNDVDFMSQSKERAT 840  
 Qy 826 TNIQPLHAQWYLGDMSECSSTCGAMORRTVECRDPSGQASATCNKALKEPDAKPCS 885  
 Db 841 TNIQPLHAQWYLGDMSECSSTCGAMORRTVECRDPSGQASATCNKALKEPDAKPCS 900  
 Qy 886 QLCPL 890  
 Db 901 QPCPL 905

## RESULT 5

US-09-321-987B-4  
 ; Sequence 4, Application US/09321987B  
 ; Patent No. US20020102210A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kimble, Judith E  
 ; APPLICANT: Bielloch, Robert H  
 ; TITLE OF INVENTION: Agent and Method for Modulating Cell Migration  
 ; FILE REFERENCE: 960296, 95386  
 ; CURRENT APPLICATION NUMBER: US/09/321, 987B  
 ; PRIOR FILING DATE: 1999-05-28  
 ; PRIOR APPLICATION NUMBER: 60/087,170  
 ; PRIOR FILING DATE: 1998-05-29  
 ; PRIOR APPLICATION NUMBER: 60/129,023  
 ; PRIOR FILING DATE: 1999-04-13  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: Patentln Ver. 2.1  
 ; SEQ ID NO 4  
 ; LENGTH: 950  
 ; TYPE: PRT  
 ; ORGANISM: Murine  
 ; US-09-321-987B-4

Query Match 48.3%; Score 2344; DB 9; Length 950;  
 Best Local Similarity 49.5%; Pred. No. 2e-177;  
 Matches 452; Conservative 143; Mismatches 241; Indels 78; Gaps 20;

Qy 14 LLLLLLLPLANGAPAPPAAGQASSELVPT--RLPG-SAGELALHLSAFKGFVRLA 70  
 Db 21 LLLASTMLLCARAGAGRPTE--EDEELVPLSLERAPGHDSITTRLRLDAFGQQLHLKIQ 78  
 Qy 71 PDDSLAPAEFKIERLGSGRATGER-----GLRGCFPSGTVNGEPESILAAVSLCRL 123  
 Db 79 PDSGFLAPGFTLTGV---GRSPSEAGHLDPTGDLAHCHYSTGVADGSAALSLCEGV 135  
 Qy 124 SSGFLDGEFEFTIQPGAGGSLAQPHRLQRMGPAPARLPREPHEVETGEG---QROE 179  
 Db 136 RGAFFYLQGEFFIOPAP-----GVATERLAPFAVEBESSARPOPHILRRRR 181  
 Qy 180 RGD-----HOEDSEESQEEAEAGASEPPPLGA-----TSRTKRFVSEA 219  
 Db 182 RGSAGAKGVWDETLPTSDSPRESQNTNRNQWVPDPTPDQAGKSPGSIKKKRFVSSP 241  
 Qy 220 RPEETLLVDAASAAAYGADLQNHILTLMSVAAATYKPSITNSINLWVAVLVYEDBKX 279  
 Db 242 RYETMLVADQSDADHSGGLKHYLLTLFSVAARPKHSITNSISLVVAVKILVYEROK 301  
 Qy 280 GPEVSDNGLLTLNFCNWQRRFNQPSDRHPEHYDTAILLTRQNFQGOGLCDTLGVADIG 339

Db 302 GPEVTSNALTLNFCNWQKQNSPSDRDPEHYDTAILFTRODLOGSH--TCDTLGMAVVG 360  
 Qy 340 TTDENKSGSVIEDBGLQAAHTLAEGLHVLMSPEHDSKPCSTRLEGWKHVNAPLPVH 399  
 Db 361 TVDPSRSQSVIEDBGLQAAHTLAEGLHVLMSPEHDSKPCSTRLEGWKHVNAPLPVH 420  
 Qy 400 LNCITLPMSPCSAMVITELLDGAGDCLDAPGAALPLPTGPGMYALYOLDOOCROIFGP 459  
 Db 421 LHSQWSPCSAMVITELLDGAGDCLDAPGAALPLPTGPGMYALYOLDOOCROIFGP 478  
 Qy 460 DPHACPNTSAQVYCAQLMC--HTDGAELPCHTRNGLSLPADGTPGPGHLCSEGSCLDEBE 518  
 Db 479 ESHGICPD--AASTCTTLMCTGTSGGLVLCQTH--FPMADSTSGEGKMCVSGKCVAKTD 534  
 Qy 519 VERPKPVNDGMAPNGMPCGSCRTGCGGVQSPHRECKPSPONGARYCLGRAYQSCHT 578  
 Db 535 MGHFATPVAGSMGPMGPMDCSRTGCGGVQYTHRECDNPFKNGKTCGSRVATRECN 594  
 Qy 579 EECPP--DGKSPREOQCEKINAYNTYMDGN--LLQWPKYAGVSPDRCKLFCRARGSE 635  
 Db 595 EDCPDNNGKTFREEQCEAHNEFSKASF--GNEPVEWETPKYAGVSPDRCKLFCRARGSE 653  
 Qy 636 FKVFENKVIDGTLGCPETLACVRGQCVKAGCHVYDSPKLDKCGVCGKNSCRKXSG 695  
 Db 654 FPLQPKVVDGTPCSPDSTSVCGQCVKAGCHVYDSPKLDKCGVCGKNSCRKXSG 713  
 Qy 696 SLTPNYGINDIVTPAGATNIDVKQSHPGVQNDGNVYALKTADQVLLNGNLAIASIE 755  
 Db 714 IVSTRPGYHDIVTPAGATNIDVKQSHPGVQNDGNVYALKTADQVLLNGNLAIASIE 773  
 Qy 756 QDILVKGITLKYSGIATLERLQSFRLPEPLVYVOLLTPGVEFPKXKTYTFVNDVDF 815  
 Db 774 QDILVKGITLKYSGIATLERLQSFRLPEPLVYVOLLTPGVEFPKXKTYTFVNDVDF 826  
 Qy 816 SMQSKERATNIIQPLHAQWYLGDMSECSSTCGAMORRTVECRDPSGQASATCNKAL 875  
 Db 827 ---KKTESFNALPTF--SEWVLEWEGESKTCGSGMQRVVCQRINDGHPASECAEV 880  
 Qy 876 KPDAKPCSQULCP 889  
 Db 881 KPDAKPCSQULCP 894

## RESULT 6

US-10-163-316-7  
 ; Sequence 7, Application US/10163316  
 ; Publication No. US20020197703A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kapeller-Liebermann, Rosana  
 ; TITLE OF INVENTION: 65552, A Human Matrix Metalloproteinase and Uses  
 ; FILE REFERENCE: MFI01-025FIRNK  
 ; CURRENT APPLICATION NUMBER: US/10/163, 316  
 ; PRIOR FILING DATE: 2002-06-05  
 ; PRIOR APPLICATION NUMBER: 60/297,863  
 ; PRIOR FILING DATE: 2001-06-13  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 7  
 ; LENGTH: 968  
 ; TYPE: PRT  
 ; ORGANISM: Mus musculus  
 ; US-10-163-316-7

Query Match 48.3%; Score 2344; DB 13; Length 968;  
 Best Local Similarity 49.5%; Pred. No. 2e-177;  
 Matches 452; Conservative 143; Mismatches 241; Indels 78; Gaps 20;

Qy 14 LLLLLLLPLANGAPAPPAAGQASSELVPT--RLPG-SAGELALHLSAFKGFVRLA 70  
 Db 38 LLLASTMLLCARAGAGRPTE--EDEELVPLSLERAPGHDSITTRLRLDAFGQQLHLKIQ 95  
 Qy 71 PDDSLAPAEFKIERLGSGRATGER-----GLRGCFPSGTVNGEPESILAAVSLCRL 123

```

Db      96 PDGSLFAPGFTLQTV---GRSPGSEAOHLDPDGLAHCFSGVNDPQSAALSLCEGV 152
Qy      124 SSGSLDGEFEFTTPOGAGGSLAQPHRLQWGPAGARPLPRGEMEVETEGS---QROE 179
Db      153 RGAFTYLOGEFTTPOAP-----GVATERLAIVAEESASAPQHILRRR 198
Qy      180 RGD-----HOEDSEESOEAEAGSEPPPLGA-----TSRTKRVESA 219
Db      199 RGSAGACGVWDEDTLPTSDSRPESQYTRNQWVPRTDPODAKPSGPGSIRKRVSSP 258
Qy      220 RFVETLVADASMAAFGADLQNHILTMGVAARIYGPISIKSINIMVYKLVIEDKX 279
Db      259 RYETMLVADOSNADFFSGSLKHYLLTFVAARFYHPIRISISLVVYKLVIEBOK 318
Qy      280 GPEVSDNGGLTLRNFQWRRFNQPSDRHPEHYDTALLTRONFCQOEGLCDTLGVADIG 339
Db      319 GPEVTSNAALTIRNFCWQKQHNPSDRDEHYDTALLFTRODLGSH--TCDTLGMADVG 377
Qy      340 TICDPNKSQSVIEDEGLQAAHTLAHELGHVLSMPHDSKPCRTLFGMGKHVMAPLFVH 399
Db      378 TVCDPSSSCSYIEDDGLQAAFTTAHELGHVFNPHDAGKQASLNGVTDGSHLMAWLS 437
Qy      400 LNOTLPMSPCSAMVLTLLDGHGDCLLDAPGALPLPTGLPGRMALYOLDQOCROI 459
Db      438 LDHSGPMSPCSAVWVTSFLDNGHGECLMDKQNPILKPSDLLG--TLYDANRQCOFTFGE 495
Qy      460 DFRHCNTSAQDVCAQIMC-HTDGAEPVLCITKNGSLPMADGTGCPGPHLCSBSCLPBEE 518
Db      496 ESKHCPD--AASTCTTLMCTGTSGLLVCCQTKH--FMAADGTSGEGKWCVSQKVNKT 551
Qy      519 VERKPYVDGGMAMPWMEGCSRTCGGVOFSHRECKDPBPONGRCLGRAPAKYOSCHT 578
Db      552 MKHATFVHSGWGMWGMWGCDSRTCGGVOYTMRECONPFPKQKXCEBKRYRYSN 611
Qy      579 EECBP-DGKSPREOCEKRYNAYNTMDGN--LIQVYKTAQVSPRDRCLFCRARGSE 635
Db      612 EDCPDNNGKTFREHCEAHNEFSKASF-GNEPVEYEMPTKTAGVSPKDRCLCTEAKIGY 670
Qy      636 FKVEEAKYIDTGLGPEPLAICVAGCCVYKAGCDHVVSPPRLDKCGVCGKNGSCRVSG 695
Db      671 FFIYQPKVVDGTPSPDSTSVCGQGVKAGCDHIDSKKFKPCGVCGGNGSCCKMGS 730
Qy      696 SLFTNRYNDIVITIPAGATNIDYKORSHPGVQNDGNVYALKTDGQYLLNGMLAISALE 755
Db      731 IYTSRTPYHDIVITIPAGATNIEVYKHNRQGRSNNNGSFLAIRADGYILLNGVFTSTLE 790
Qy      756 QDIIVKGTILKYSGSIATLERLQSRPLPEPLTVQLLTVPGEVFPKVKYKTFPVNDVF 815
Db      791 QDLTKGTVLRYSGSSALERIRSFPLKEPLTIQVLMV--GHALRPKIKFYFV----- 843
Qy      816 SMQSKKATNTNIQPLHQAQVLDGMEGCSSTCGAGMORRYTECRPDSGQASATCNKAL 875
Db      844 ----KKKTESNAPLPTF--SEWVIEHMEGCSCKTCGSGMQRVYVCGIINGHPSACEKVE 897
Qy      876 KPEDAKPCESQLCP 889
Db      898 KPASTRPCADLPCP 911

```

```

RESULT 7
US-09-373-658-2
; Sequence 2, Application US/09373658
; Publication No. US20030092900A1
; GENERAL INFORMATION:
; APPLICANT: Inuela-Ariape, Luisa
; APPLICANT: Hastings, Gregg A.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Uonak, Zdenka L.
; APPLICANT: Trull, Stephen H.
; APPLICANT: Fromwald, James A.
; APPLICANT: Terrett, Jonathan A.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides

```

```

; FILE REFERENCE: 1488.1070006
; CURRENT APPLICATION NUMBER: US/09/373, 658
; CURRENT FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 950
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-373-658-2

Query Match
Best Local Similarity 49.7%; Pred. No. 3,4e-16;
Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;

Qy      11 LPFLILLILLIPLAGRPARPAAGQASSELVVP--TFLPSAGELALHSAPGKGYLR 68
Db      16 VFTLLLLAALLAAS-DLGRPSE--EDDELVPELEKAPGH-GTTRLRLLAFDQQLDLE 71
Qy      69 LAPDSFLAPEKIERLQSGSRATGG-----ERGLRGCFSGTYNGEPESLAASLCRGL 123
Db      72 LRPDSSFLAPFTLQNV---GRKSGSERPLPETDLAHCFYAGTVGDDSSAALSLCEGV 128
Qy      124 SSGSLDGEFEFTTPOGAGGSLAQPHRLQWGPAGARPLP-----RGPEMEV-----ET 172
Db      129 RGAFTYLOGEFTTPOAP-----GVATERLAIVAEESASAPQHILRRRQGVGTGCV 183
Qy      173 GEGORGEHDHDEDESEOEAEAGASEPP-----PPLGATS--RTKRVSEARFV 222
Db      184 VDDERPRPKATEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE 243
Qy      223 ETLVADASMAAFYADLQNHILTMGVAARIYKHPISIKSINIMVYKLVIEDKMBE 282
Db      244 ETLVADOSMAEFHSGSLKHYLLTFVAARLYKHPISIRNSVLYVYKLVIEBOKPE 303
Qy      283 VSDNGGLTLRNFQWRRFNQPSDRHPEHYDTALLTRONFCQOEGLCDTLGVADIGTIC 342
Db      304 VTSNALTIRNFCWQKQHNPSDRDEHYDTALLFTRODLGSH--TCDTLGMADVGTVIC 362
Qy      343 DENKCSVIEDDGLQAAHTLAHELGHVLSMPHDSKPCRTLFGMGKHVMAPLFVHLNQ 402
Db      363 DSRSCSVIEDDGLQAAFTTAHELGHVFNPHDAGKQASLNGVTDGSHLMAWLSNDH 422
Qy      403 TLPMSPCSAMVLTLLDGHGDCLLDAPGALPLPTGLPGRMALYOLDQOCROI 462
Db      423 SQPMSPCSAVWVTSFLDNGHGECLMDKQNPILKPSDLLG--TLYDANRQCOFTFGE 480
Qy      463 HCPNTSAQDVCAQIMC-HTDGAEPVLCITKNGSLPMADGTGCPGPHLCSBSCLPBEEVER 521
Db      481 HCPD--AASTCTTLMCTGTSGLLVCCQTKH--FMAADGTSGEGKWCVSQKVNKTDRGH 536
Qy      522 PXPVVDGGMAMPWMEGCSRTCGGVOFSHRECKDPBPONGRCLGRAPAKYOSCHTBE 581
Db      537 PTFPHSGWGMWGMWGCDSRTCGGVOYTMRECONPFPKQKXCEBKRYRYSNCLTEC 596
Qy      582 PP-DGKSPREOCEKRYNAYNTDM--DGNLQVYKTAQVSPRDRCLFCRARGSEFKYV 639
Db      597 PDNNGKTFREHCEAHNEFSKASF-GNEPVEYEMPTKTAGVSPKDRCLCTEAKIGY 656
Qy      640 EAKYIDTGLGPEPLAICVAGCCVYKAGCDHVVSPPRLDKCGVCGKNGSCRVSGSLTP 699
Db      657 QPKVVDGTPSPDSTSVCGQGVKAGCDHIDSKKFKPCGVCGGNGSCCKKISGSVTS 716
Qy      700 TNYGVNDIVITIPAGATNIDYKORSHPGVQNDGNVYALKTDGQYLLNGMLAISALEQ 759
Db      717 AKPGYHDITIPAGATNIEVYKHNRQGRSNNNGSFLAIRADGYILLNGVFTSTLEQ 776
Qy      760 VKGTILKYSGSIATLERLQSRPLPEPLTVQLLTVPGEVFPKVKYKTFPVNDVPSMS 819
Db      777 YKGVVLRYSGSSALERIRSFPLKEPLTIQVLMV--GHALRPKIKFYFV----- 825
Qy      820 SKERATNTNIQPLHQAQVLDGMEGCSSTCGAGMORRYTECRPDSGQASATCNKALPED 879

```



Fri May 7 12:18:02 2004

us-09-989-687-4\_1.rapb

Page 6

Db 826 KKKKSEFNAL-PTFSA-WIIEWGECSEKSCELGMQRRLVECDINDINGOPASECAKEVPAAS 883  
QY 880 AKPCESQLCP 889  
Db 884 TRPCADHPCP 893

RESULT 8

US-09-989-687-2  
; Sequence 2, Application US/09989667  
; Publication No. US20040002449A1  
; GENERAL INFORMATION:  
; APPLICANT: Hastings, Gregg A.  
; APPLICANT: Ruben, Steven M.  
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides  
; FILE REFERENCE: 1488.107000D  
; CURRENT APPLICATION NUMBER: US/09/989,687  
; CURRENT FILING DATE: 2001-11-21  
; NUMBER OF SEQ ID NOS: 126  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 950  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-989-687-2

Query Match 48.0%; Score 2328.5; DB 11; Length 950;  
Best Local Similarity 49.7%; Pred. No. 3,4e-176;  
Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;

QY 11 LPTLLILLILLPLLAGAPAPPAAGQASEIYVP--TRLPGSAGELALHLSAFKGFVLR 68  
Db 16 VPTLLILLAALLAVS-DALGRPSE--EDEELVPELERAPGH-GTTRLRLHAFDQQLDLE 71  
QY 69 LAPDGSFLAPEFKIRLGGSGRATG-----ERGLRGCFPGTGVNGEPESLAASLRCGL 123  
Db 72 LRPDSSFLAPGFTLQNV--GRKSGSETPLPETDLAHCFYGTGVNGDSSAALSLCEGV 128  
QY 124 GGSFLIDGEBEFTIOPQAGAGSLAQPHRLQRMGPAGARPLP-----RGPEWEV---ET 172  
Db 129 RGAFTILGSAVFIQPLPAAS-----ERLATAFGKAPPAIQLPHLRNRQGVGTGV 183  
QY 173 GEGORGERGDHODESEESQEEAEAGASEPP-----PPLGATS-RTRKRVSEARFV 222  
Db 184 VDDEPRPTGKATEDEDEGEDEGEPQMSPODPALQGVGPTGTGSIKKRRFVSSHRYV 243  
QY 223 ETLIVADAMAAYGADLQNHILTLMSVAARIYKHSIKNSINLMVYVILVDEBKQPE 282  
Db 244 ETLIVADQSMAEFHSGLKHVLLTLFVAARLYKHSIRNSVSLVAVKILVIHDEQKPE 303  
QY 283 VSDNGGLTRNFCNMQRRFNQPSDRHPHEHYDTAILLTRONFCQOEGLCDTLGVADIGTIC 342  
Db 304 VTSNALLTRNFCNMQKQNPSPDRDAEHYDTAILFTRDLCGSG--TCDTLGMADVGTVC 362  
QY 343 DPKKGSVTEDEGLQAHTLAHELGVLSMHPDSDKPCRRLLFGPMGKHVVAAPLFEVHNO 402  
Db 363 DPKRSQSVTEDEGLQAHTLAHELGVFMHPDDAQOCASLNGVADQSHMAASLMSNIDH 422  
QY 403 TLPWSPCSAMVLTLELDGSHGDCILDAFGAALPLPTGLFGRMALYQLDQOCHQIFGPDR 462  
Db 423 SQWMSCSAYMTISFLDNGHGBGLMDKPNPIQLPDDLPG--TSYDARQCQFTGEDSK 480  
QY 463 HCPNTSAQDVCAQJLC-HTDGAEPFLCHTKNGSLPNAADGTCPPGPHCSGSGCLPEBEER 521  
Db 481 HCPD--AASTCSTLWGTGSGVLYCQPKH--FPMADGTSCEBGMKINGACVNTKDKH 536  
QY 522 PKXVVDGWAAPMGWMECSEKRTCGGAVGFHSRECKDPEPONGRXYCLGRBAKYQSCHTEBC 581  
Db 537 FDIPIFGSNMGMPKMDCCRTGCGGVQYIMRECDNVFPGNGGVYGGKRVYRSQNLBDC 596  
QY 582 PP-DGKSFREOCEKTNAYITDM-DGNLLQNVPKYAGVSPDRCLLFRAGGRSEFKYF 639  
Db 597 PDNNGKTFREEOCEANEFESKASFGSGPAVEMIPKYAGVSPKDRCLLCOAGKIGYFFVL 656

QY 640 EAKYIDGTLGSEPTALICVRGQCYAKGCDHVVDSPRKLDKGCVGCKGNSCRKYSGLTP 699  
Db 657 QPKVVDGTPCSPDSTVCGQCVAVAGCDRIIDSKKFPDKGCVGCKGNSCTCKKISGTS 716  
QY 700 TNYGVNDIYTPAGATNIDVKORSHPGVQNDGNVYALKTADGOYLLNGLNLAISAEODIL 759  
Db 717 AKPGYHDITITPTGATNIEVKORNGSGRRNSFLATAAGTYILNDYTLSTIEODIM 776  
QY 760 VKGTILKYSGSTATTERRLOSFRPLPEPLTVQLLITPGEVFPKXKTYTFVVDVDFSMQS 819  
Db 777 YKGVVLYRSGSSAALERIRSFPLEPLTIQVLTV-GNALRKRKYTFV----- 825  
QY 820 SKERATNTIOPLLHAQWLTGDMSECSSTCGAGMQRRTVECDRPSGQASATCNKALKEPD 879  
Db 826 KKKKSEFNAL-PTFSA-WIIEWGECSEKSCELGMQRRLVECDINDINGOPASECAKEVPAAS 883  
QY 880 AKPCESQLCP 889  
Db 884 TRPCADHPCP 893

RESULT 9  
US-10-105-929-2  
; Sequence 2, Application US/10105929  
; Publication No. US20020137142A1  
; GENERAL INFORMATION:  
; APPLICANT: Holtzman, Douglas D.J.  
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83  
; FILE REFERENCE: 09404/041001  
; CURRENT APPLICATION NUMBER: US/10/105,929  
; PRIOR FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/130,491  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-07  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,108  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-05  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/054,961  
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 967  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-105-929-2

Query Match 48.0%; Score 2328.5; DB 13; Length 967;  
Best Local Similarity 49.7%; Pred. No. 3,4e-176;  
Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;

QY 11 LPTLLILLILLPLLAGAPAPPAAGQASEIYVP--TRLPGSAGELALHLSAFKGFVLR 68  
Db 33 VPTLLILLAALLAVS-DALGRPSE--EDEELVPELERAPGH-GTTRLRLHAFDQQLDLE 88  
QY 69 LAPDGSFLAPEFKIRLGGSGRATG-----ERGLRGCFPGTGVNGEPESLAASLRCGL 123  
Db 89 LRPDSSFLAPGFTLQNV--GRKSGSETPLPETDLAHCFYGTGVNGDSSAALSLCEGV 145  
QY 124 GGSFLIDGEBEFTIOPQAGAGSLAQPHRLQRMGAARPLP-----RGPEWEV---ET 172  
Db 146 RGAFTILGSAVFIQPLPAAS-----ERLATAFGKAPPAIQLPHLRNRQGVGTGV 200  
QY 173 GEGORGERGDHODESEESQEEAEAGASEPP-----PPLGATS-RTRKRVSEARFV 222  
Db 201 VDDEPRPTGKATEDEDEGEDEGEPQMSPODPALQGVGPTGTGSIKKRRFVSSHRYV 260  
QY 223 ETLIVADAMAAYGADLQNHILTLMSVAARIYKHSIKNSINLMVYVILVDEBKQPE 282  
Db 261 ETLIVADQSMAEFHSGLKHVLLTLFVAARLYKHSIRNSVSLVAVKILVIHDEQKPE 320  
QY 283 VSDNGGLTRNFCNMQRRFNQPSDRHPHEHYDTAILLTRONFCQOEGLCDTLGVADIGTIC 342

Fri May 7 12:18:02 2004

us-09-989-687-4\_1.rapb

Page 7

321 VTSNAALTRNFCWQKXHPSPSDRDAEHYDIALFTRODLCSQ--TCDTLGMADVITVC 379  
343 DPNKSCSVIEDEGLQAAHTLAHEIGHVLSMPHDSKPCRTLFGPMKHHVMAPLFVHLNQ 402  
380 DPNKSCSVIEDEGLQAAFTTAHEIGHVFNMPHDDAKOCASLNGNODSHMMAISLNDH 439  
403 TLPMSPCSAMVLTLLDGGHDCILDAFGALPLPTGLPGMALYQDDQCRQIFGDFR 462  
440 SQPMSPCSAMVLTSLDNGHSCILMDKQNPQLPGLPG--TSYANRQCGTFGSDSK 497  
463 HCPNTSAQDVCAQIMC-HTDGAEPILCHTKNGSLPMWADGTPCGPHGLCSGSCLEEVER 521  
498 HCPD--AASCTSLMTCTSGGLVLCQTKH--FPWADGTSCEGKWCINCKCVNKTDRKH 553  
522 PKPVVDGMAWPMWMBEGSRTCCGGVQFSPHRECKDEPONGGRYCLGRARAKYSCHTREC 581  
554 FDTFFHSGMWGMWPMWDCSRTCCGGVQYTMRECDNVPKNGGKICBGRVRYRSCNLEDC 613  
582 PP-DGKSFREOQCEKXNANYTDM-DGNLLQWPKYAGVSPRDCXLCFRARGRSEPKVF 639  
614 PDNNGKTFREOCEAHNBFSAKSGSPAVEMIPKXAGVSPKRCXLCQAKIGYFVL 673  
640 EAYVIGTLGPEPTLAICVRGQCYKAGCDHVDSPRKLDKCGVCGGKNSCRKYSGLTP 699  
674 QPKVVDGTPCSPDSTSVCCVQGCYKAGCDRIIDSKKFKDKCGVCGGNGSTCKKISGVS 733  
700 TNYGNDIVITIPAGATNIDVKORSHPGVQNDGNALKTADGGYLLNGNLASAIEDIL 759  
734 AKPVHDIITITGATNIEVQKORNGRNNGSFLAIKAADGTYIINADTLSTLEBDM 793  
760 VKGTLKXSGSIATLERLQSFRLPEPLTVOLLTVPEGEVPPKXKTYTFVYVNDVDSMOS 819  
794 YKGVLYRSGSSAALERISFSPLEPLTIQVLTIV-GNALRPKIKYVYFV----- 842  
820 SKERTNTIILQPLHMQVLDGWSGSCSTCGAGMORRTVECRDPSGGAATCNKALPDE 879  
843 KKKESFPAI-PTBSA-WVIEBWGECSSKCELGMRRLVECRDINGOPASCAKEVFPAS 900  
880 AKPCESQUCP 889  
901 TRPCADHPCP 910  
RESULT 10  
US-10-115-286-2  
Sequence 2, Application US/10115286  
Publication No. US2003016065A1  
GENERAL INFORMATION:  
APPLICANT: Jonak, Zdenka  
Trull, Stephen  
Formwald, James  
Teirre, Jonathan  
Hasting, Gregg  
TITLE OF INVENTION: No. US2003016065A1 Integrin ligand ITGL-TSP  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Ratner & Prestia  
STREET: Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/115,286  
FILING DATE: 04-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/845,496

FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GH-70000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0701  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 967 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-115-286-2  
Query Match 48.0%; Score 2328.5; DB 14; Length 967;  
Best Local Similarity 49.7%; Pred. No. 3,4e-176;  
Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;  
11 LPELILLLLLLPLRAGAPAPAGGASSELVVP--TRLPGSAGELALHLSAFKGFYLR 68  
33 VPTLLALLAALLAVS-DALGRPE--EDELVVPELERAPGH-GTTRRLHAFDQDLLE 88  
69 LAPDSSFLAPEFKIERLIGSGSRATGQ----ERGLRGCFSGGTGNGEPELSLAVALSICRGL 123  
89 LRPDSSFLAPGFTLQNV--GRKSGSEPLPEPTDLAHCFYSGTVNGDPSSAALSLICEGV 145  
124 SGSFLIDGEEFTIIOQAGGSLAQPHRLQRMKPARAPLP-----RGPEMEV---ET 172  
146 RGAFFYLGAAYFIQPPAAS-----ERLATAAGKEXPAPLOFHLRRRQGVGTGCV 200  
173 GEGORQERGDHODESEESCEEAEGASBP-----PLIGATS-RTKSFVSEARFV 222  
201 VDDEPRPTGALEDEDEDETEGEDEGPQWSPODPALQGVQGTGNGSIKKKFFVSHRYV 260  
223 ETLIVADASMAAFYGDILQNHILITLMSVAARIYKBPISKNSINLMVYKLYEDEKMGSE 282  
261 ETLIVADQSMAEFHSGGLKHYLITLPSVAARYKBPISKNSINSLVYKLVHDEKMGSE 320  
283 VSDNGGLTLRNFQNMORRFFNOPSDRHEHYDIALITRONFGQEBGLCDTLGADIGTIC 342  
321 VTSNAALTRNFCWQKXHPSPSDRDAEHYDIALFTRODLCSQ--TCDTLGMADVITVC 379  
343 DPNKSCSVIEDEGLQAAHTLAHEIGHVLSMPHDSKPCRTLFGPMKHHVMAPLFVHLNQ 402  
380 DPNKSCSVIEDEGLQAAFTTAHEIGHVFNMPHDDAKOCASLNGNODSHMMAISLNDH 439  
403 TLPMSPCSAMVLTLLDGGHDCILDAFGALPLPTGLPGMALYQDDQCRQIFGDFR 462  
440 SQPMSPCSAMVLTSLDNGHSCILMDKQNPQLPGLPG--TSYANRQCGTFGSDSK 497  
463 HCPNTSAQDVCAQIMC-HTDGAEPILCHTKNGSLPMWADGTPCGPHGLCSGSCLEEVER 521  
498 HCPD--AASCTSLMTCTSGGLVLCQTKH--FPWADGTSCEGKWCINCKCVNKTDRKH 553  
522 PKPVVDGMAWPMWMBEGSRTCCGGVQFSPHRECKDEPONGGRYCLGRARAKYSCHTREC 581  
554 FDTFFHSGMWGMWPMWDCSRTCCGGVQYTMRECDNVPKNGGKICBGRVRYRSCNLEDC 613  
582 PP-DGKSFREOQCEKXNANYTDM-DGNLLQWPKYAGVSPRDCXLCFRARGRSEPKVF 639  
614 PDNNGKTFREOCEAHNBFSAKSGSPAVEMIPKXAGVSPKRCXLCQAKIGYFVL 673  
640 EAYVIGTLGPEPTLAICVRGQCYKAGCDHVDSPRKLDKCGVCGGKNSCRKYSGLTP 699  
674 QPKVVDGTPCSPDSTSVCCVQGCYKAGCDRIIDSKKFKDKCGVCGGNGSTCKKISGVS 733  
700 TNYGNDIVITIPAGATNIDVKORSHPGVQNDGNALKTADGGYLLNGNLASAIEDIL 759

```

Db 734 AKPGYHDIITITPGATNIEVKQNRQGRNRNGSFLAIKAADGTLYLNGDYTLSTLEBDIM 793
QY 760 VKGTLIKSGSIATLERLQSFRLPEPLTVOLLTVPEVPPPKYKTFEFPVNDVDFSMOS 819
Db 794 YKGVVLRYSGSSALERIRSFPLKEPLTIQVLTIV-GNALRPKIKYTFV----- 842
QY 820 SKERATNIIQPLHAQWVLGDMWSESSSTCGAGMQRRTVECRDPSGQASATCNKALPED 879
Db 843 KKKKESFNAI-PTFSA-WVIEEMGECSSKCSLGMQRRLVECRDINGQDPASCAKAVKPAS 900
QY 880 AKPCESQQLCP 889
Db 901 TRPCADHPCP 910

RESULT 11
US-09-373-658-125
; Sequence 125, Application US/09373658
; Publication No. US20030092900A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A.
; APPLICANT: Iruela-Arispe, Luisa
; APPLICANT: Ruben, Steven M.
; APPLICANT: Jonak, Zdenka L.
; APPLICANT: Trull, Stephen H.
; APPLICANT: Fromwald, James A.
; APPLICANT: Terrett, Jonathan A.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.1070006
; CURRENT APPLICATION NUMBER: US/09/373,658
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 125
; LENGTH: 968
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-373-658-125

Query Match 48.0%; Score 2328.5; DB 10; Length 968;
Best Local Similarity 49.7%; Pred. No. 3.4e-176;
Matches 452; Conservative 146; Mismatches 249; Indels 63; Gaps 22;

QY 11 LPEFLILLPLLAGAPAPAPAGQASSELVVP--TRLPGSAGELALHLAFAFGKGVFLR 68
Db 34 VPLILLALLAALAVS-DALGRPS--EDELYVPELEBARGH-GTTRLRIHAFOQDLLE 89
QY 69 LAPDDSLAPPEFKIERLGGSGRATG---ERGLRGCFSGTVNGEPESLAASVLCRGL 123
Db 90 LRDDSSFLAPGFTLQNV--GRKSGSETPLPETDLACHFSGTVNGDPSAALSLCEGV 146
QY 124 SSGFLLDGEFTIOPGAGGSLAQPHRLQRMGPAGARPLP-----RGPEWEV---ET 172
Db 147 RGAFFYLLGAYFLOPLPAS-----ERLATAPEKXPAPLOPHILRRRQDVGTCGV 201
QY 173 GEGORQERPDHDESEESOEERAGASBP-----PPLGATS-RTKRFVSEARFV 222
Db 202 VDDEPRPTGAETREDEDEGEDEGPQMSPOBALQGVGQGTGSGIKKRFVSSHXYV 261
QY 223 ETLVLVADAGAAATYAGDLQNHITLMSVAARIYKHSINSONINMVMVKYLIEDKMPRE 282
Db 262 ETLVLVADAGAAATYAGDLQNHITLMSVAARIYKHSINSONINMVMVKYLIEDKMPRE 321
QY 283 VSDNGGLTLANFCNMQRRFNQPSDRAPHEHYDTAILLTRONFGQEGELCDTLGVADIGTIC 342
Db 322 VTISNALITLRFNCKQKQHPSPDRADENHDTAILLTRQDLGSGQ-TCTTLGMADVGYVC 380
QY 343 DPKKSGSVTDEGLQAHTLAHLGYLSNPHDDSPCTRLLRGPMPKHHVMAPLFHLNQ 402
Db 381 DPKRSGSVTDEGLQAHTLAHLGYLSNPHDDSPCTRLLRGPMPKHHVMAPLFHLNQ 440
QY 403 TLMSPGSAVYLTLLDGGHGDCLDAPGAALPLPTGLPGRNALYOLDOQCQOIFGPDR 462

```

```

Db 441 SQPMSPCSAMYTSTFLNGHGECLMDKPQNPLOLFDLPG--TSVDANKQOFTGEDSK 498
QY 463 HCPNTSAQDVCAQLWC-HTDGAELPCHTRNGSLPWADGTPCGPHLCSEBGLPEEYVER 521
Db 499 HCPD--NASTGSLWCTGTSGLVLCQTKH--FPAADGTSCEBGMKCNKGKVNKTRDKH 554
QY 522 PKPVVDGMAWPMWPMGECSRTGCGGVQFSHRECHDPEFQNGRXYCLGRARAKQSHTEC 581
Db 555 PDLTFHSGVMWGMWPMGDGSRCTGCGGVQYMERCDNPVFNKGKYEGRKVRYSNLEDC 614
QY 582 PP-DGKSFREOQCEKNANAYTDM--DGNLLQWVPXYAGVSPERCKLFCRAGRSEFYVF 639
Db 615 PDNNGKTFREBOCEANESKASPSGSAVEMIPKXAGVSPERCKLFCQAKGIGYFVL 674
QY 640 EAKVIDGLCGPELTALCYRGQCVTAQCDHVVDSFRKLDKCGVCGGKNSCKXVSGSLTP 699
Db 675 QPKVVDGTPCSDPSTSVCGQCVXAGCDRIIDSKKEDKCGVCGANSTCKKISGVTS 734
QY 700 TNYGVNDIYITPAGATNIDVKORSHPGVQNDGNVYALATKADQYLLNGALISAIEDIL 759
Db 735 AKPGYHDIITITPGATNIEVKQNRQGRNRNGSFLAIKAADGTLYLNGDYTLSTLEBDIM 794
QY 760 VKGTLIKSGSIATLERLQSFRLPEPLTVOLLTVPEVPPPKYKTFEFPVNDVDFSMOS 819
Db 795 YKGVVLRYSGSSALERIRSFPLKEPLTIQVLTIV-GNALRPKIKYTFV----- 843
QY 820 SKERATNIIQPLHAQWVLGDMWSESSSTCGAGMQRRTVECRDPSGQASATCNKALPED 879
Db 844 KKKKESFNAI-PTFSA-WVIEEMGECSSKCSLGMQRRLVECRDINGQDPASCAKAVKPAS 901
QY 880 AKPCESQQLCP 889
Db 902 TRPCADHPCP 911

```

```

RESULT 12
US-09-741-151-4
; Sequence 4, Application US/09741151
; Publication No. US20020086400A1
; GENERAL INFORMATION:
; APPLICANT: ZHU, Shiaoqing et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OR INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CLO01.005
; CURRENT APPLICATION NUMBER: US/09/741,151
; CURRENT FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Human
US-09-741-151-4

Query Match 48.0%; Score 2328; DB 12; Length 931;
Best Local Similarity 49.8%; Pred. No. 3.6e-176;
Matches 451; Conservative 143; Mismatches 250; Indels 62; Gaps 21;

QY 15 LILLILLPLLAGAPAPAPAGQASSELVVP--TRLPGSAGELALHLAFAFGKGVFLRIAPD 72
Db 1 LILLIAALLAVSDALGRPS--EDELYVPELEBARGH-GTTRLRIHAFOQDLLEIRPD 57
QY 73 DGLAPPEFKIERLGGSGRATG---ERGLRGCFSGTVNGEPESLAASVLCRGLSGSF 127
Db 58 SSFLAPPEFTLQNV--GRKSGSETPLPETDLACHFSGTVNGDPSAALSLCEGVARGAF 114
QY 128 LLDGEFTIOPGAGGSLAQPHRLQRMGPAGARPLP-----RGPEWEV---ETGEGQ 176
Db 115 YLDEAFYLOPLPAS-----ERLATAPEKXPAPLOPHILRRRQDVGTCGVVDDE 169
QY 177 ROERGHQDESEESOEERAGASBP-----PPLGATS-RTKRFVSEARFVETTL 226

```

```

Db 170 PRETKAETDEDEGEDEGEDEGAQWSPDPAQGVGQPTGTGSIKRRFVSSHRYETML 229
Qy 227 VAPASMAAFYADLQNHILTLMSVAARIYKGPSIKNSINLMVYKVLIVEDEKMPPEVSDN 286
Db 230 VAOQSAEFHSGSLKHYLLTLFSVAALYKHPISIRNSVSLVYVKIIVIHDEQGPATVN 289
Qy 287 GGLTFNFCNWRFPNOPSDRHPEHYDTAILLTRQNFQGEGLCDTLGVADIGTICDPNK 346
Db 290 AALLTRFCNWKQKHPNPSDRDAEHYDTAILLTRQDLCGSGQ--TCDTLGMADVGVCDPSR 348
Qy 347 SCSTIEBEGLOAHTLAHELGHVLSMPHDSKPCITLFGPMGKHVNAFLFVHNLCTLM 406
Db 349 SCSTIEBEGLOAHTLAHELGHVFNHDDAKQASLNGVQDSHMASSLNLSQHM 408
Qy 407 SPCSAMYTELLDGHDCCLIDAFGAALPLPTGLFGPMALYOLDQOCROIFGDPFHCPN 466
Db 409 SPCGAMYTSLFDNGHGECLMDKQNPQLPGLDLPF--TSYDANRQOQFTFGEDSKHCD 466
Qy 467 TSAQVCAQLWC-HTDGAEPCLTKRNSLPMADGTPCGPHLCSGSLCPBEVEVERPXV 525
Db 467 --AASTSTLMCTGTSGVLVCCQTKH--FPMADGTSGBGKWCINGCNVAKTRKHPDTP 522
Qy 526 VDGWMAFMPWGECSRTCCGGVOPSHRECKDPEPONGARYCLGRAPAYQCHTECP--D 584
Db 523 FHSGWGMWGPWGDSCRTCCGGVQYTMRECDNPVPKNGKTCBGKRVYRSQNLDECPDN 582
Qy 585 GKSPREOQCEKYNANYTDM--DGNLQWPKYAGVSPRDKCLFCAPARGSEFKVPEAKY 643
Db 583 GKTPREOQCEAHNFSSKASFGSGPAVEMIPKYAGVSPDKCLCOAKGIGYFVLQPKY 642
Qy 644 IDGLQCPETALIVRGQCVKAGCDHVDSPRKLDKCGVCGKNSCRKYSGLTPTNYG 703
Db 643 VDGPPCPDSTSVCCQCVKAGCDRIIDSKKFKDKCGVCGNGSTCKKISSGVTAKRG 702
Qy 704 YNDVITIPAGATNIDVYKORSHPVQNDGNVLAALKTADGQYLLNGNLASIEBODILVKG 763
Db 703 YNDITIPGATNIEVQNRQGRNNGSFLAIAADGTIYLNGDYTLSTLEODIMYKGV 762
Qy 764 ILKSGSATIERLOSRPLPEPTVOLLTVPGVFPKXKTFEVPNDVDFSMQSKER 823
Db 763 VLKSGSAALEIRISFSLPEPLTIQVLTV--GNALRPKIKYTFV-----KKKK 811
Qy 824 ATTNIQPLHAQWVLGWSGSSCTCGAGQRTYECRDPGQASATCNKALXPEDAKPC 883
Db 812 ESFNAI-PTFSA-WVIEWGECSKSCJELGMQRRLVECRDINGQASCAEVDPASTRFC 869
Qy 884 ESQJCP 889
Db 870 ADHPCP 875

RESULT 13
US-09-989-687-126
; Sequence 126, Application US/09989687
; Publication No. US2004000249A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.107000D
; CURRENT APPLICATION NUMBER: US/09/989,687
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 126
; LENGTH: 967
; TYPE: PRT
; ORGANISM: ITGL-TSP
; US-09-989-687-126

Query Match 47.8%; Score 2320.5; DB 11; Length 967;
Best Local Similarity 49.5%; Pred. No. 1.5e-175;
Matches 450; Conservative 147; Mismatches 250; Indels 63; Gaps 22;

```

```

Qy 11 LPFLILLILLPLARGAPARPAAGQASLVP--TRLPSSAGELALHLSAFGKGVLR 68
Db 33 VPTLLLAALLAAS--DALGRPSE--EDEELVPELEKRVPH--GTRRLRLAFLQOQDLD 88
Qy 69 LAPDSEFLAEPKIERLOGSGRATG-----ERGLRGCFSGTNGEESLAASVSLGRL 123
Db 89 VPPDSEFLAEPKIQNV--GRKSGSTPLPETLACHFTYSGTNGDSSAAALSLCEGV 145
Qy 124 SGSPLLDSEFTTPOGAGGSLAQPHRLQRMWPAARLP-----RGPEWY----ET 172
Db 146 RGAFYLIGEAFTIQPLPAAS-----ERLATAPEKPEPAFLQHLRRNRQGDVGATCGV 200
Qy 173 GEGQKQERQDQDESEESQEEBAEAGSEPP-----PPLATLS--RTKRFVSEARFV 222
Db 201 VDDEPRFTGKATEDEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGEDEGE 260
Qy 223 ETLVADASMAAFYADLQNHILTLMSVAARIYKGPSIKNSINLMVYKVLIVEDEKMP 282
Db 261 ETLVADASMAEFHSGSLKHYLLTLFSVAALYKHPISIRNSVSLVYVKIIVIHDEQGP 320
Qy 283 VSDNGGLTKRNFQWRRFPNOPSDRHPEHYDTAILLTRQNFQGEGLCDTLGVADIGTIC 342
Db 321 VTSNAALTLRFPQWQKHPNPSDRDAEHYDTAILLTRQDLCGSGQ--TCDTLGMADVGV 379
Qy 343 DENKSCSYIEBEGLOAHTLAHELGHVLSMPHDSKPCITLFGPMGKHVNAFLFVHNLQ 402
Db 380 DSRSCSYIEBEGLOAHTLAHELGHVFNHDDAKQASLNGVQDSHMASSLNLSH 439
Qy 403 TLPWSPCSAMYTELLDGHDCCLIDAFGAALPLPTGLFGPMALYOLDQOCROIFGDP 462
Db 440 SQPMFPCSGYMITSLDNGHGECLMDKQNPQLPGLDLPF--TSYDANRQOQFTFGEDSK 497
Qy 463 HCPNPSADQVCAQLWC-HTDGAEPCLTKRNSLPMADGTPCGPHLCSGSLCPBEVEER 521
Db 498 HCPD--AASTSTLMCTGTSGVLVCCQTKH--FPMADGTSGBGKWCINGCNVAKTRK 553
Qy 522 PKFVVDGMAFMPWGECSRTCCGGVOPSHRECKDPEPONGARYCLGRAPAYQCHTEC 581
Db 554 FDTPFHSGWGMWGPWGDSCRTCCGGVQYTMRECDNPVPKNGKTCBGKRVYRSQNLDC 613
Qy 582 PP-DKSPREOQCEKYNANYTDM--DGNLQWPKYAGVSPRDKCLFCAPARGSEFKV 639
Db 614 PNNKGTREBQCEAHNFSSKASFGSGPAVEMIPKYAGVSPDKCLCOAKGIGYFVL 673
Qy 640 EAKVIDGTLQCPETALIVRGQCVKAGCDHVDSPRKLDKCGVCGKNSCRKYSGLTPT 699
Db 674 QPKYVDGTPCSPDSTSVCCQCVKAGCDRIIDSKKFKDKCGVCGNGSTCKKISSGVT 733
Qy 700 TNYGNDVITIPAGATNIDVYKORSHPVQNDGNVLAALKTADGQYLLNGNLASIEB 759
Db 734 AKPGYNDITIPGATNIEVQNRQGRNNGSFLAIAADGTIYLNGDYTLSTLEODIM 793
Qy 760 VKGTIKSGSATIERLOSRPLPEPTVOLLTVPGVFPKXKTFEVPNDVDFSMQSK 819
Db 794 YKGVLRISGSSAALEIRISFSLPEPLTIQVLTV--GNALRPKIKYTFV----- 842
Qy 820 SKERATTNIQPLHAQWVLGWSGSSCTCGAGQRTYECRDPGQASATCNKALXPED 879
Db 843 KKKESFNAL-PTFSA-WVIEWGECSKSCJELGMQRRLVECRDINGQASCAEVDPAS 900
Qy 880 AKPCESQJCP 889
Db 901 TRPCADHPCP 910

RESULT 14
US-09-802-582-8
; Sequence 8, Application US/09802582
; Publication No. US20020086354A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.

```

```

; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: 0734-323001
; CURRENT APPLICATION NUMBER: US 09/802,582
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 09/128,709
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: US 60/054,645
; PRIOR FILING DATE: 1997-08-04
; PRIOR APPLICATION NUMBER: US 09/130,491
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: US 60/054,966
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 09/388,280
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 09/388,279
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-802-582-8

```

```

Query Match      46.9%; Score 2274.5; DB 12; Length 481;
Best Local Similarity 82.1%; Pred. No. 2.5e-172;
Matches 403; Conservative 35; Mismatches 40; Indels 13; Gaps 2;

```

```

QY 401 NOTLPMSCSAAYLTLELDGSHGDCILDAPGALPLPTGLPGRMALYOLDQCCQIFGPD 460
DB 3 NKTLPWSPCSAYLTLELDGSHGDCILD-----DGHSTLYELDDQCCQIFGPD 50
QY 461 FRHCNTSAQDYCAQLWC-HTDGAFLCHTNGSLPMADGTPCGPGLHCSBSCCLPEEV 519
DB 51 FRHCNTSVEIDCYGLWCRHRSDDEPCHTNGASLMDGTPCGPGLHCLDGS CVLREEV 110
QY 520 EEPKPVVDGMAFMGPWGECSRTCCGGVQFSHRECKDEPONGRYCLGRAXYOSCHTE 579
DB 111 ENPKAVVDGDMFMPWGPWQCSRTCCGGIQFSNRECDNAPONGRFLCGERVKYOSCKTE 170
QY 580 ECPDGSFRBQCCCKNAVNTTMDGNLLQWVPKYAGVSPDRCKLFCRARGSEFKVF 639
DB 171 ECPDGSFRBQCCCKNAVNTTMDGNLLQWVPKYAGVSPDRCKLFCRARGSEFKVF 230
QY 640 EAKVIDGTLGPEETLAI CVRGQCVAGCDHVVDSPRKLDKCGVCGGKNSCRKXVGSILTP 699
DB 231 ETKVIDGTLGPEETLAI CVRGQCVAGCDHVVDSPRKLDKCGVCGGKNSCRKXVGSILTP 290
QY 700 TNYGNDIVITIPAGATNIDVKORSHPGVQNDGNVYALKTADQYLLNGNLAI SAIBDIL 759
DB 291 PSYGNDIVITIPAGATNIDVKORSHPGVQNDGSYALKTANGQYLLNGNLAI SAIBDIL 350
QY 760 VKGTLIKYSGSIATLERLOSFRPLPEPLTVOLLTVGSEVFPKXKTFEVPNDVDFSMOS 819
DB 351 MKGTLIKYSGSIATLERLOSFRPLPEPLTVOLLTVGSEVFPKXKTFEVPNDVDFSMOS 410
QY 820 SKERATNTIIQPLLAQWYLGDMSECSSTCGAGMORRYTECRDPGQASATCNKALKPED 879
DB 411 SKERASTNTIIQSLPYAEWVLGDMSECPSTCGGMRRYTECRDPGQASATCNKALKPED 470
QY 880 AKPCESQCLPL 890
DB 471 AKPCGSQPCLL 481

```

```

RESULT 15
US-10-105-929-8
; Sequence 8, Application US/10105929

```

```

; Publication No. US20020137142A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US 10/105,929
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/130,491
; PRIOR FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/054,961
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Rattus rattus
; US-10-105-929-8

```

```

Query Match      46.9%; Score 2274.5; DB 13; Length 481;
Best Local Similarity 82.1%; Pred. No. 2.5e-172;
Matches 403; Conservative 35; Mismatches 40; Indels 13; Gaps 2;

```

```

QY 401 NOTLPMSCSAAYLTLELDGSHGDCILDAPGALPLPTGLPGRMALYOLDQCCQIFGPD 460
DB 3 NKTLPWSPCSAYLTLELDGSHGDCILD-----DGHSTLYELDDQCCQIFGPD 50
QY 461 FRHCNTSAQDYCAQLWC-HTDGAFLCHTNGSLPMADGTPCGPGLHCSBSCCLPEEV 519
DB 51 FRHCNTSVEIDCYGLWCRHRSDDEPCHTNGASLMDGTPCGPGLHCLDGS CVLREEV 110
QY 520 EEPKPVVDGMAFMGPWGECSRTCCGGVQFSHRECKDEPONGRYCLGRAXYOSCHTE 579
DB 111 ENPKAVVDGDMFMPWGPWQCSRTCCGGIQFSNRECDNAPONGRFLCGERVKYOSCKTE 170
QY 580 ECPDGSFRBQCCCKNAVNTTMDGNLLQWVPKYAGVSPDRCKLFCRARGSEFKVF 639
DB 171 ECPDGSFRBQCCCKNAVNTTMDGNLLQWVPKYAGVSPDRCKLFCRARGSEFKVF 230
QY 640 EAKVIDGTLGPEETLAI CVRGQCVAGCDHVVDSPRKLDKCGVCGGKNSCRKXVGSILTP 699
DB 231 ETKVIDGTLGPEETLAI CVRGQCVAGCDHVVDSPRKLDKCGVCGGKNSCRKXVGSILTP 290
QY 700 TNYGNDIVITIPAGATNIDVKORSHPGVQNDGNVYALKTADQYLLNGNLAI SAIBDIL 759
DB 291 PSYGNDIVITIPAGATNIDVKORSHPGVQNDGSYALKTANGQYLLNGNLAI SAIBDIL 350
QY 760 VKGTLIKYSGSIATLERLOSFRPLPEPLTVOLLTVGSEVFPKXKTFEVPNDVDFSMOS 819
DB 351 MKGTLIKYSGSIATLERLOSFRPLPEPLTVOLLTVGSEVFPKXKTFEVPNDVDFSMOS 410
QY 820 SKERATNTIIQPLLAQWYLGDMSECSSTCGAGMORRYTECRDPGQASATCNKALKPED 879
DB 411 SKERASTNTIIQSLPYAEWVLGDMSECPSTCGGMRRYTECRDPGQASATCNKALKPED 470
QY 880 AKPCESQCLPL 890
DB 471 AKPCGSQPCLL 481

```

```

Search completed: May 7, 2004, 11:57:19
Job time : 58 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:56:21 ; Search time 26 Seconds  
(without alignments)  
3292.710 Million cell updates/sec

Title: US-09-989-687-4

Perfect score: 890

Sequence: 1 MFPAAPAPRWLPULLLLLL.....CNALAKPEDAKPCESQCLPL 890

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

1: PIR.\*  
2: PIR1.\*  
3: PIR2.\*  
4: PIR3.\*  
5: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1.5	2165	2	T21371	hypothetical prote
2	1.3	205	2	S37804	hypothetical prote
3	1.3	837	2	T00355	hypothetical prote
4	1.2	951	2	T00017	gene ADAMTS-1, prot
5	1.2	263	2	SE7346	interleukin 15 rec
6	1.1	655	1	A46688	hepatocyte growth
7	1.1	231	2	A49265	FLT3/Flt-2 ligand
8	1.1	331	2	T15587	hypothetical prote
9	1.1	480	2	A31589	carboxypeptidase C
10	1.1	491	2	JC6197	stromelysin 3 (BC
11	1.1	492	2	A44399	stromelysin 3 (BC
12	1.1	550	2	T47158	hypothetical prote
13	1.1	1015	2	JC5263	transmembrane tyro
14	1.1	1015	2	JC5062	phogrin precursor
15	1.1	1277	2	T30532	neural cell adhesi
16	1.1	112	2	A60111	sporozoite antigen
17	1.0	133	2	SE7038	probable membrane
18	1.0	134	2	A40959	secretin precursor
19	1.0	144	2	T18867	hypothetical prote
20	1.0	171	2	S15502	hypothetical prote
21	1.0	182	2	JC7236	receptor-activativ
22	1.0	185	2	A49957	CD45-associated 30
23	1.0	189	2	JC7262	receptor activativ
24	1.0	197	2	SE1372	membrane protein 1
25	1.0	206	2	A55412	lymphocyte phospho
26	1.0	244	2	S18948	centrocyclin precu
27	1.0	285	2	JC5639	1-acylglycerol-3-p
28	1.0	348	2	G02297	gene N33 protein -
29	1.0	368	2	AEO086	filmbrial protein I

30	9	1.0	381	2	S35940	class I histocompa
31	9	1.0	382	2	B84527	hypothetical prote
32	9	1.0	384	2	S25771	gaal protein - mou
33	9	1.0	402	1	BOHUS	sex steroid-bindin
34	9	1.0	409	2	H70303	ubiquinol-cytochro
35	9	1.0	429	2	T17215	hypothetical prote
36	9	1.0	452	2	F71327	probable signal re
37	9	1.0	481	2	T22406	hypothetical prote
38	9	1.0	500	2	SE6522	cartilage matrix p
39	9	1.0	557	2	S12359	nicotinic acetylch
40	9	1.0	602	2	SE9198	prostoglandin G/H
41	9	1.0	602	2	S39782	cyclooxigenase 1 -
42	9	1.0	626	1	NBHUIA	platelet glycoprot
43	9	1.0	871	2	S47518	cadherin - African
44	9	1.0	941	2	TYWVMD	protein-tyrosine k
45	9	1.0	964	2	T15746	hypothetical prote

#### ALIGNMENTS

RESULT 1  
T21371  
hypothetical protein F25H8.3 - Caenorhabditis elegans  
C/Spectrum: Caenorhabditis elegans  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C/Accession: T21371, T24896  
R/Gadastly, S.  
submitted to the EMBL Data Library, February 1996  
A/Reference number: Z19413  
A/Accession: T21371  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-2165 <M12>  
A/Cross-references: EMBL:Z69360; PIDD:CAA93287.1; GSPDB:GN00022; CESP:F25H8.3  
A/Experimental source: clone F25H8  
R/Gadastly, S.  
submitted to the EMBL Data Library, February 1996  
A/Reference number: Z19949  
A/Accession: T24896  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-2165 <M12>  
A/Cross-references: EMBL:Z69361; PIDD:CAA93288.1; GSPDB:GN00022; CESP:F25H8.3  
A/Experimental source: clone T13H10  
C/Genetics:  
A/Gene: CESP:F25H8.3  
A/Map position: 4  
A/Intons: 31/1; 52/1; 135/2; 193/3; 216/1; 266/1; 495/2; 547/3; 584/3; 634/2; 744/1; 8

Query Match  
Best Local Similarity 100.0%; Pred. No. 0.00083;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 MGECSTCGGQV 548  
Db 612 MGECSTCGGQV 624

RESULT 2  
S37804  
hypothetical protein YKL47C - yeast (Saccharomyces cerevisiae)  
N/Alternate names: hypothetical protein YKL601  
C/Spectrum: Saccharomyces cerevisiae  
C/Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 19-Apr-2002  
A/Accession: S37804; S37977; S44581  
R/Vandenberg, M.; Boile, P.; Dion, C.; Porteballe, D.; Hilger, F.  
submitted to the EMBL Data Library, September 1993  
A/Description: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LAF4 1  
A/Reference number: S37786  
A/Accession: S37804  
A/Molecule type: DNA  
A/Residues: 1-205 <YAN>

A/Cross-references: EMBL:Z26877; NID:9407482; PIDN:CAA81507.1; PID:9407501  
 A/Experimental source: Strain S288C  
 A/Vanderbol, M.; Bolle, P.A.; Dion, C.; Portetalle, D.; Hliger, F.  
 submitted to the Protein Sequence Database, March 1994  
 A/Reference number: S37976  
 A/Accession: S37977  
 A/Molecule type: DNA  
 A/Residues: 1-205 <VA2>  
 A/Cross-references: EMBL:Z28146; NID:9486250; PIDN:CAA81987.1; PID:9486251; MIPS:YKL147C  
 A/Experimental source: Strain S288C  
 A/Vanderbol, M.; Bolle, P.A.; Dion, C.; Portetalle, D.; Hliger, F.  
 Yeast 10, 35-40, 1994  
 A/Title: DNA sequencing of a 36.2 kb fragment located between the FAS1 and LMP4 loci of  
 A/Reference number: S44563  
 A/Accession: S44581  
 A/Status: nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A/Residues: 1-205 <VA3>  
 A/Cross-references: EMBL:Z26877; NID:9407482; PIDN:CAA81507.1; PID:9407501  
 A/Experimental source: Strain S288C  
 A/Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1993  
 C/Genetics:  
 A/Cross-references: SGD:S0001630  
 A/Map position: 11L  
 C/Superfamily: Saccharomyces hypothetical protein YKL147C

Query Match 1.3%; Score 12; DB 2; Length 205;  
 Best Local Similarity 100.0%; Pred. No. 0.0012;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 13 FLLILLILPL 24  
 DB 73 FLLILLILPL 84

RESULT 3  
 T00355  
 Hypothetical protein KIAA0688 - human  
 C/Species: Homo sapiens (man)  
 C/Date: 01-Feb-1999 #sequence\_revision 01-Feb-1999 #text\_change 21-Jul-2000  
 C/Accession: T00355  
 R:Shikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.  
 DNA Res. 5, 169-176, 1998  
 A/Title: Prediction of the coding sequences of unidentified human genes. X. The complete  
 A/Reference number: Z14142; MUID:98403880; PMID:9734811  
 A/Accession: T00355  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-837 <ISH>  
 A/Cross-references: EMBL:AB014588; NID:93327189; PIDN:BAA3163.1; PID:93327190  
 A/Experimental source: brain  
 C/Genetics:  
 A/Genes: KIAA0688  
 C/Superfamily: thrombospondin type 1 repeat homology <THR3>  
 F/519-575/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 1.3%; Score 12; DB 2; Length 837;  
 Best Local Similarity 100.0%; Pred. No. 0.0038;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 539 CSRTGGGVQPS 550  
 DB 532 CSRTGGGVQPS 543

RESULT 4  
 T00017  
 gene ADAMTS-1 protein - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 21-Jul-2000  
 C/Accession: T00017  
 R:Kono, K.; Iizasa, H.; Ohno, S.; Matsushima, K.  
 Genomics 46, 466-471, 1997

A/Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 gen  
 A/Reference number: Z14055; MUID:9810583; PMID:9441751  
 A/Accession: T00017  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-951 <KUN>  
 A/Cross-references: EMBL:AB001735; NID:92809056; PIDN:BAA24501.1; PID:92809057  
 A/Experimental source: strain 129SVJ  
 C/Genetics:  
 A/Genes: ADAMTS-1  
 A/Introns: 228/1; 343/3; 388/1; 444/1; 539/3; 602/1; 660/3; 719/2  
 C/Superfamily: thrombospondin type 1 repeat homology  
 F/542-598/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 1.3%; Score 12; DB 2; Length 951;  
 Best Local Similarity 100.0%; Pred. No. 0.0042;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 706 DIVTIPAGATNI 717  
 DB 724 DIVTIPAGATNI 735

RESULT 5  
 S57346  
 Interleukin 15 receptor precursor - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 05-Nov-1999  
 C/Accession: S57346  
 R:Giri, U.G.; Kumaki, S.; Ahdieh, M.; Friend, D.C.; Loomis, A.; Shanebeck, K.; Dubose,  
 EMO J. 14, 3654-3663, 1995  
 A/Title: Identification and cloning of a novel IL-15 binding protein that is structural  
 A/Reference number: S57346; MUID:95369237; PMID:7641685  
 A/Accession: S57346  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-263 <GIR>  
 A/Cross-references: EMBL:U22339; NID:9551105; PIDN:AA052240.1; PID:9551106  
 C/Superfamily: complement factor H repeat homology  
 C/Keywords: cytokine receptor  
 F/36-96/Domain: complement factor H repeat homology <FH2>

Query Match 1.2%; Score 11; DB 2; Length 263;  
 Best Local Similarity 100.0%; Pred. No. 0.015;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 14 LLLILLILPL 24  
 DB 18 LLLILLILPL 28

RESULT 6  
 A46688  
 hepatocyte growth factor activator (EC 3.4.21.-) precursor [validated] - human  
 C/Species: Homo sapiens (man)  
 C/Date: 21-Sep-1993 #sequence\_revision 25-Aug-1995 #text\_change 08-Dec-2000  
 C/Accession: A46688  
 R:Miyazawa, K.; Shimomura, T.; Kitamura, A.; Kondo, J.; Morimoto, Y.; Kitamura, N.  
 J. Biol. Chem. 268, 10024-10028, 1993  
 A/Title: Molecular cloning and sequence analysis of the cDNA for a human serine protea  
 d coagulation factor XII.  
 A/Reference number: A46688; MUID:93252878; PMID:7683665  
 A/Accession: A46688  
 A/Molecule type: mRNA  
 A/Residues: 1-655 <MTY>  
 A/Cross-references: DBJ:D14012; NID:9219680; PIDN:BAA03113.1; PID:9219681  
 A/Experimental source: liver (mRNA); serum (protein)  
 A/Note: sequence extracted from NCBI backbone (NCBI:J31227, NCBI:P.131228)  
 C/Genetics:  
 A/Genes: GDB:HGFAC; HGFA; HGFAP  
 A/Cross-references: GDB:9954514  
 A/Map position: 4p16-4p16

C:Function: activates hepatocyte growth factor by specific proteolytic cleavage  
 A:Pathway: tissue repair and regeneration  
 C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homology;  
 C:Keywords: glycoprotein; hydrolase; Kringle; liver; plasma; serine proteinase  
 F:1-34/Domain: signal sequence status predicted <Sig>  
 F:108-148/Domain: fibronectin type II repeat homology <LF2>  
 F:164-197/Domain: EGF homology <EG1>  
 F:202-237/Domain: fibronectin type I repeat homology <LF1>  
 F:243-278/Domain: EGF homology <EG2>  
 F:286-367/Domain: Kringle homology <KRG>  
 F:373-407/Domain: hepatocyte growth factor activator light chain status experimental  
 F:408-655/Domain: hepatocyte growth factor activator heavy chain status experimental  
 F:408-641/Domain: trypsin homology <TRY>  
 F:440-48,290,468,492,546/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:164-175,169-186,188-197,202-230,228-237,245-256,250-267,269-278,286-367,307-349,338-36  
 F:447,497,598/Active site: His, Asp, Ser #status predicted

Query Match 1.2%; Score 11; DB 1; Length 655;  
 Best Local Similarity 100.0%; Pred. No. 0.031;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 PFLLLLLLLL 22  
 DB 18 PFLLLLLLLL 28

RESULT 7  
 A:Accession: A49265  
 C:Species: Mus musculus (house mouse)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 08-Oct-1999  
 C:Accession: A49265; 1493467; 1493467; 583290  
 R:Lyman, S.D.; James, L.; Vanden Bos, T.; de Vries, P.; Brasel, K.; Gliniak, B.; Holling  
 D.; Williams, D.E.; Beckmann, M.P.  
 Cell 73, 1157-1167, 1993  
 A:Title: Molecular cloning of a ligand for the flt3/flk-2 tyrosine kinase receptor: a p  
 A:Reference number: A49265; MUID:94084791; PMID:7505204  
 A:Accession: A49265  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-221 <LTM>  
 A:Cross-references: GB:U23636; NID:g439441; PID:AAA93436.1; PID:g439442  
 R:Lyman, S.D.; Stocking, K.; Davison, B.; Fletcher, F.; Johnson, L.; Escobar, S.  
 Oncogene 11, 1165-1172, 1995  
 A:Title: Structural analysis of human and murine flt3 ligand genomic loci.  
 A:Reference number: 139075; MUID:96032581; PMID:7566977  
 A:Accession: 149347  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-163, 'G', 165, 'HAG' <RES>  
 A:Cross-references: EMBL:U29875; NID:g1072039; PID:AAA90952.1; PID:g1072041  
 A:Accession: 149346  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-197, 'L', 198-231 <RES>  
 A:Cross-references: EMBL:U29875; NID:g1072039; PID:AAA90951.1; PID:g1072040  
 R:Hamam, C.; Culpepper, J.; Campbell, D.; McClanahan, T.; Zurawski, S.; Bazan, J.F.; Ka  
 felt, A.; Muench, M.; Kelnier, G.; Nankiawa, R.; Kemick, D.; Roncetto, M.G.; Zlotnik, A  
 Nature 368, 643-648, 1994  
 A:Title: Ligand for FLT3/Flk2 receptor tyrosine kinase regulates growth of haematopoietic  
 A:Reference number: S43290; MUID:94195428; PMID:8145851  
 A:Accession: S43290  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-197, 'L', 198-231 <HAN>  
 A:Experimental source: clone T10  
 A:Note: the sequence from Fig. 2c is inconsistent with that from Fig. 2a in having 4-Val  
 C:Genetics: 11/3; 49/3; 67/3; 115/3; 164/1; 224/3  
 C:Keywords: transmembrane protein

Query Match 1.1%; Score 10; DB 2; Length 231;

Best Local Similarity 100.0%; Pred. No. 0.13;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 LLLLLLLPL 24  
 DB 190 LLLLLLLPL 199

RESULT 8  
 A:Accession: T15587  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T15587  
 R:Favell, T.  
 submitted to the EMBL Data Library, November 1995  
 A:Description: The sequence of C. elegans cosmid C24A3.  
 A:Reference number: Z18373  
 A:Accession: T15587  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-331 <FAV>  
 A:Cross-references: EMBL:U40424; NID:g1065542; PID:g1065546; PID:AAA81458.1; CESP:C24A  
 C:Genetics: 3  
 A:Gene: CESP:C24A3.3  
 A:Introns: 27/1; 73/2; 136/1; 193/2

Query Match 1.1%; Score 10; DB 2; Length 331;  
 Best Local Similarity 100.0%; Pred. No. 0.18;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 LLLLLLLPL 23  
 DB 320 LLLLLLLPL 329

RESULT 9  
 A:Accession: A31589  
 C:Species: Homo sapiens (man)  
 C:Date: 11-May-1989 #sequence\_revision 11-May-1989 #text\_change 05-May-2000  
 C:Accession: A31589  
 R:Gallagher, N.J.; Gilleman, N.; Harris, A.; van der Horst, G.T.J.; Verheijen, F.W.; Gal  
 Cell 54, 755-764, 1988  
 A:Title: Expression of cDNA encoding the human "protective protein" associated with lys  
 A:Reference number: A31589; MUID:88311078; PMID:3136930  
 A:Accession: A31589  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-480 <GAL>  
 A:Cross-references: GB:M22960; GB:J03159; GB:M18453; NID:g190282; PID:AAA36476.1; PID:  
 C:Superfamily: serine carboxypeptidase  
 C:Keywords: hydrolase; serine carboxypeptidase  
 F:176,400,457/Active site: Ser, Asp, His #status predicted

Query Match 1.1%; Score 10; DB 2; Length 480;  
 Best Local Similarity 100.0%; Pred. No. 0.25;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 PFLLLLLLLL 22  
 DB 10 PFLLLLLLLL 19

RESULT 10  
 JCG197  
 etromelysin 3 (EC 3.4.24.-) - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 11-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 17-Mar-1999  
 C:Accession: JCG197  
 R:Okada, A.; Saez, S.; Miumi, Y.; Bassett, P.  
 Gene 185, 187-193, 1997



A>Title: Rat stromelysin 3: cDNA cloning from healing skin wound, activation by furin at  
 A/Reference number: J06197; MUID:97208872; PMID:9055814  
 A/Contents: Skin wounds  
 A/Accession: J06197  
 A/Molecule type: mRNA  
 A/Residues: 1-491 <OKA>  
 A/Cross-references: GB:U46034  
 C/Comment: This protein is a member of the matrix metalloproteinase family.  
 C/Superfamily: Interstitial collagenase; hemopexin repeat homology; matrix metalloprote  
 C/Keywords: hydrolase; metalloproteinase; zinc; zymogen  
 F:52-261/Domain: matrix metalloproteinase homology <MMP>  
 F:291-483/Domain: hemopexin repeat homology <PXN>  
 F:184,218,222,228/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
 F:218,222,228/Binding site: zinc, catalytic (His) (active) #status predicted  
 F:219/Active site: Glu #status predicted

Query Match 1.1%; Score 10; DB 2; Length 491;  
 Best Local Similarity 100.0%; Pred. No. 0.25;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLP 23  
 Db 21 LLLLLLLLP 30

RESULT 11  
 A43399  
 stromelysin 3 (EC 3.4.24.-) - mouse  
 N/Alternate names: matrix metalloproteinase 11 (MMP11)  
 C/Species: Mus musculus (house mouse)  
 C/Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 10-Jul-1998  
 C/Accession: A44399  
 R/RefSeq: O.; Wolf, C.; Linacher, J.M.; Huttin, P.; Wendling, C.; Lemeur, M.; Baeset, F  
 U: Cell Biol. 119, 997-1002, 1992  
 A>Title: The breast cancer-associated stromelysin-3 gene is expressed during mouse mamma  
 A/Reference number: A44399; MUID:93054930; PMID:1423645  
 A/Accession: A44399  
 A/Status: preliminary; not compared with conceptual translation  
 A/Molecule type: nucleic acid  
 A/Residues: 1-492 <LEF>  
 A/Note: sequence extracted from NCBI backbone (NCBIP:117216)  
 C/Superfamily: Interstitial collagenase; hemopexin repeat homology; matrix metalloprote  
 C/Keywords: hydrolase; metalloproteinase; zinc; zymogen  
 F:52-262/Domain: matrix metalloproteinase homology <MMP>  
 F:292-484/Domain: hemopexin repeat homology <PXN>  
 F:184,219,223,229/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
 F:219,223,229/Binding site: zinc, catalytic (His) (active) #status predicted  
 F:220/Active site: Glu #status predicted

Query Match 1.1%; Score 10; DB 2; Length 492;  
 Best Local Similarity 100.0%; Pred. No. 0.25;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLP 23  
 Db 21 LLLLLLLLP 30

RESULT 12  
 T47158  
 hypothetical protein DKFZP762C1110.1 - human (fragment)  
 C/Species: Homo sapiens (man)  
 C/Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C/Accession: T47158  
 R/Blum, H.; Bauersachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.  
 submitted to the Protein Sequence Database, March 2000  
 A/Reference number: Z24379  
 A/Accession: T47158  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-550 <AAA>  
 A/Cross-references: EMBL:AL162080  
 A/Experimental source: adult melanoma (W60 cell line); clone DKFZP762C1110

C/Genetics:  
 A/Note: DKFZP762C1110.1

Query Match 1.1%; Score 10; DB 2; Length 550;  
 Best Local Similarity 100.0%; Pred. No. 0.28;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 CSRTCGGCGV 548  
 Db 154 CSRTCGGCGV 163

RESULT 13  
 JC5263  
 transmembrane tyrosine phosphatase-like protein, ICAR - human  
 C/Species: Homo sapiens (man)  
 C/Date: 25-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 20-Jun-2000  
 C/Accession: JC5263  
 R/Smith, P.D.; Barker, K.T.; Wang, J.; Lu, Y.J.; Shipley, J.; Crompton, M.R.  
 Biochem. Biophys. Res. Commun. 229, 402-411, 1996  
 A>Title: ICAR, a novel member of a new family of transmembrane, tyrosine phosphatase-1  
 A/Reference number: JC5263; MUID:97127415; PMID:8954911  
 A/Accession: JC5263  
 A/Status: nucleic acid sequence not shown  
 A/Molecule type: mRNA  
 A/Residues: 1-1015 <SMI>  
 A/Cross-references: GB:Y08569; NID:G1644377; PID:CAA69880.1; PID:G1644378  
 C/Comment: This protein has an intracellular protein tyrosine phosphatase like protein.  
 C/Superfamily: protein-tyrosine-phosphatase, receptor type N; protein-tyrosine-phosphat  
 C/Keywords: phosphoprotein  
 F:770-994/Domain: protein-tyrosine-phosphatase homology <PTP>  
 F:945/Active site: Cys (phosphocysteine intermediate) #status predicted  
 F:951/Binding site: substrate phosphate (Arg) #status predicted

Query Match 1.1%; Score 10; DB 2; Length 1015;  
 Best Local Similarity 100.0%; Pred. No. 0.47;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLLP 23  
 Db 7 LLLLLLLLP 16

RESULT 14  
 JC5062  
 phogrin precursor - human  
 N/Contents: protein-tyrosine-phosphatase (EC 3.1.3.48) receptor type  
 C/Species: Homo sapiens (man)  
 C/Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 21-Jun-2002  
 C/Accession: JC5062; T46903  
 R/Kawasaki, E.; Hutton, J.C.; Bisenbath, G.S.  
 Biochem. Biophys. Res. Commun. 227, 440-447, 1996  
 A>Title: Molecular cloning and characterization of the human transmembrane protein tyr  
 A/Reference number: JC5062; MUID:97032784; PMID:8878534  
 A/Accession: JC5062  
 A/Contents: 1set  
 A/Molecule type: mRNA  
 A/Residues: 1-1015 <RAW>  
 A/Cross-references: GB:U66702; NID:G1620663; PID:AA50742.1; PID:G1620664  
 R/Smith, P.D.; Barker, K.T.; Wang, J.; Lu, Y.J.; Shipley, J.; Crompton, M.R.  
 Biochem. Biophys. Res. Commun. 229, 402-411, 1996  
 A>Title: ICAR, a novel member of a new family of transmembrane, tyrosine phosphatase-1  
 A/Reference number: JC5263; MUID:97127415; PMID:8954911  
 A/Accession: JC5263  
 A/Status: nucleic acid sequence not shown  
 A/Molecule type: mRNA  
 A/Residues: 1-207, 'S', 209-246, 'G', 248-322, 'R', 324, 'N', 326-1015 <SMI>  
 A/Cross-references: GB:Y08569; NID:G1644377; PID:CAA69880.1; PID:G1644378  
 R/Ansorge, W.; Winkler, U.; Mewes, H.W.; Weil, B.; Wiemann, S.  
 submitted to the Protein Sequence Database, February 2000  
 A/Reference number: Z24134  
 A/Accession: T46903  
 A/Status: preliminary

A:Molecule type: mRNA  
 A:Residues: 'DA',714,771-1015 <AAA>  
 A:Cross-references: EMBL:AL157451  
 A:Experimental source: adult amygdala; clone DKFZp761A0712  
 C:Comment: This protein has an intracellular protein tyrosine phosphatase like protein.  
 C:Genetics:  
 A>Note: DKFZp761A0712.1  
 C:Superfamily: protein-tyrosine-phosphatase, receptor type N; protein-tyrosine-phosphatase  
 C:Keywords: phosphoprotein; phosphoric monoester hydrolase; transmembrane protein; tyrosine  
 F:1-17/Domain: signal sequence #status predicted <SIG>  
 F:18-1015/Product: phogrin #status predicted <MAT>  
 F:615-638/Domain: transmembrane #status predicted <TM>  
 F:770-994/Domain: protein-tyrosine-phosphatase homology <PTP2>  
 F:945/Active site: Cys (phosphocysteine intermediate) #status predicted  
 F:951/Binding site: substrate phosphate (Arg) #status predicted

Query Match 1.1%; Score 10; DB 2; Length 1015;  
 Best Local Similarity 100.0%; Pred. No. 0.47;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 LLLLLLLP 23  
 |||||  
 DB 7 LLLLLLLP 16

RESULT 15  
 T30532  
 neural cell adhesion molecule L1 homolog - Fugu rubripes  
 C:Species: Fugu rubripes  
 C:Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 01-Mar-2002  
 C:Accession: T30532  
 R:Ripoldi, Tunicliffe, G.R.; Pltzer, M.; Nyakatura, G.; Elgar, G.S.; Brenner, S.; Rosen  
 Submitted to the EMBL Data Library, September 1997  
 A:Description: Analysis of the genomic loci of Fugu rubripes homologs of the human discoidin domain receptor  
 A:Reference number: Z20848  
 A:Accession: T30532  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1277 <RIB>  
 A:Cross-references: EMBL:AF026198, NID:G3098263, PID:G3098264, PIDN:AACT5580.1  
 C:Genetics:  
 A:introns: 42/1; 47/1; 81/2; 149/1; 190/1; 247/1; 285/2; 347/1; 391/1; 440/1; 477/2; 531/2  
 A>Note: L1-CAM  
 C:Superfamily: neural cell adhesion molecule L1; fibronectin type III repeat homology;

Query Match 1.1%; Score 10; DB 2; Length 1277;  
 Best Local Similarity 100.0%; Pred. No. 0.57;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 LLLLLLLP 25  
 |||||  
 DB 19 LLLLLLLP 28

Search completed: May 7, 2004, 12:00:48  
 Job time : 28 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 12:00:12 ; Search time 54 Seconds  
(without alignments)  
4574.711 Million cell updates/sec

Title: US-09-989-687-4

Perfect score: 890  
Sequence: 1 MFPAAPAPRWLPFLILL...CNKALXPDAPKESQLCPL 890

Scoring table: CLIGO  
Gapop 60.0, Gapext 60.0

Searched: 1140673 seqs, 27756755 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1140673

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 45 summaries

Database:

Published Applications AA:\*  
1: /cgn2\_6/prodata/2/pubpa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/prodata/2/pubpa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/prodata/2/pubpa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/prodata/2/pubpa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/prodata/2/pubpa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/prodata/2/pubpa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/prodata/2/pubpa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/prodata/2/pubpa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/prodata/2/pubpa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/prodata/2/pubpa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/prodata/2/pubpa/US09\_NEW\_PUB.pep:\*  
12: /cgn2\_6/prodata/2/pubpa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/prodata/2/pubpa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/prodata/2/pubpa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/prodata/2/pubpa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/prodata/2/pubpa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/prodata/2/pubpa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/prodata/2/pubpa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	890	100.0	890	10	US-09-373-658-4
2	890	100.0	890	11	US-09-989-687-4
3	710	79.8	924	12	US-10-425-114-39107
4	263	29.6	364	9	US-09-764-903-57
5	217	24.4	245	9	US-09-918-171A-11
6	35	3.9	905	9	US-09-918-171A-11
7	30	3.4	481	12	US-09-802-582-8
8	30	3.4	481	13	US-10-105-929-8
9	30	3.4	481	14	US-10-365-227-8
10	13	1.5	2150	9	US-09-321-878-2
11	13	1.5	2150	9	US-09-800-729-155
12	12	1.3	438	9	US-09-963-791-22
13	12	1.3	438	12	US-10-419-276-22
14	12	1.3	518	9	US-09-803-589-10
15	12	1.3	551	12	US-09-802-582-16

16	12	1.3	551	13	US-10-105-929-16	Sequence 16, Appl
17	12	1.3	551	14	US-10-365-227-16	Sequence 16, Appl
18	12	1.3	575	12	US-10-358-283-12	Sequence 12, Appl
19	12	1.3	589	9	US-09-963-791-12	Sequence 12, Appl
20	12	1.3	589	12	US-10-419-276-12	Sequence 12, Appl
21	12	1.3	727	9	US-09-445-023A-12	Sequence 12, Appl
22	12	1.3	727	14	US-10-097-597-12	Sequence 12, Appl
23	12	1.3	727	14	US-10-097-580-12	Sequence 12, Appl
24	12	1.3	757	9	US-09-963-791-12	Sequence 24, Appl
25	12	1.3	757	12	US-10-419-276-24	Sequence 24, Appl
26	12	1.3	837	10	US-09-946-374-317	Sequence 317, App
27	12	1.3	837	12	US-10-206-915-352	Sequence 352, App
28	12	1.3	837	12	US-10-196-670-352	Sequence 352, App
29	12	1.3	837	12	US-10-201-858-352	Sequence 352, App
30	12	1.3	837	12	US-10-205-890-352	Sequence 352, App
31	12	1.3	837	12	US-10-208-024-352	Sequence 352, App
32	12	1.3	837	12	US-10-201-853-352	Sequence 352, App
33	12	1.3	837	12	US-10-358-283-11	Sequence 11, Appl
34	12	1.3	837	12	US-10-174-581-352	Sequence 352, App
35	12	1.3	837	12	US-10-176-483-352	Sequence 352, App
36	12	1.3	837	12	US-10-176-749-352	Sequence 352, App
37	12	1.3	837	12	US-10-176-914-352	Sequence 352, App
38	12	1.3	837	12	US-10-176-915-352	Sequence 352, App
39	12	1.3	837	12	US-10-065-485A-317	Sequence 317, App
40	12	1.3	837	12	US-10-013-907A-317	Sequence 317, App
41	12	1.3	837	12	US-10-015-499A-317	Sequence 317, App
42	12	1.3	837	12	US-10-176-484-352	Sequence 352, App
43	12	1.3	837	12	US-10-180-550-352	Sequence 352, App
44	12	1.3	837	12	US-10-183-014-352	Sequence 352, App
45	12	1.3	837	12	US-10-187-738-352	Sequence 352, App

#### ALIGNMENTS

RESULT 1  
US-09-373-658-4  
Sequence 4, Application US/09373658  
Publication No. US20030092900A1  
GENERAL INFORMATION:  
APPLICANT: Ituela-Artepe, Luisa  
APPLICANT: Hastings, Gregg A.  
APPLICANT: Ruben, Steven M.  
APPLICANT: Jonak, Zdenek L.  
APPLICANT: Trull, Stephen H.  
APPLICANT: Fromwald, James A.  
TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides  
FILE REFERENCE: 1488.107006  
CURRENT APPLICATION NUMBER: US/09/373,658  
CURRENT FILING DATE: 1999-08-13  
NUMBER OF SEQ ID NOS: 125  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 4  
LENGTH: 890  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-373-658-4  
Query Match 100.0%; Score 890; DB 10; Length 890;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MFPAAPAPRWLPFLILL...LPARCAPAPAGAGASLIVPTRLPGSAGELATHLSA 60  
DB 1 MFPAAPAPRWLPFLILL...LPARCAPAPAGAGASLIVPTRLPGSAGELATHLSA 60  
QY 61 FKGKGVFLAPDDSLAPDFKIERLGGSGRATGGERGIRGCFBSGTYNGPESLAASVLC 120  
DB 61 FKGKGVFLAPDDSLAPDFKIERLGGSGRATGGERGIRGCFBSGTYNGPESLAASVLC 120  
QY 121 RGLSGFLIDDEERTIOQAGAGSLAOPHRICRWGPAGARLPFGPEWEVETGQROER 180  
DB 121 RGLSGFLIDDEERTIOQAGAGSLAOPHRICRWGPAGARLPFGPEWEVETGQROER 180

```

Db      121 RGLSGSFLDGEFFTIOPQAGAGSLAOPHRLQRMWGPAGARELPGRPEMEVETGEGORER 180
Qy      181 GDHODESEESQEEBAEGASEPPPLGATSRTRKFEVSEARFVETLLVADASMAAFYADL 240
Db      181 GDHODESEESQEEBAEGASEPPPLGATSRTRKFEVSEARFVETLLVADASMAAFYADL 240
Qy      241 QNHILITMSVAARIYKHPISIKNSINLMVYKYLIVDEKMGPEVSDNGGLTLRNFQNMQR 300
Db      241 QNHILITMSVAARIYKHPISIKNSINLMVYKYLIVDEKMGPEVSDNGGLTLRNFQNMQR 300
Qy      301 FNQPSDRPHBYDITALLTRQNFQCGEGICDTLGVADIGTICDPNKSQSVIIEDEGLQAAH 360
Db      301 FNQPSDRPHBYDITALLTRQNFQCGEGICDTLGVADIGTICDPNKSQSVIIEDEGLQAAH 360
Qy      361 TLAEHLGHVLSMPHDSKPCRTLFGPMGKHVMAPLFVHLNQTLLPMSPCSAWYLTLLDG 420
Db      361 TLAEHLGHVLSMPHDSKPCRTLFGPMGKHVMAPLFVHLNQTLLPMSPCSAWYLTLLDG 420
Qy      421 GHGDCILDPAGALPLPTGLPGRMALYOLDQCCROI FGPDRHCNPTSAQDVCAQLMCHT 480
Db      421 GHGDCILDPAGALPLPTGLPGRMALYOLDQCCROI FGPDRHCNPTSAQDVCAQLMCHT 480
Qy      481 DGAEPICHTKNGSLPMAOGTFCGPHLCSGSCLPPEEVERPKPVVDGMAPMGWBGCS 540
Db      481 DGAEPICHTKNGSLPMAOGTFCGPHLCSGSCLPPEEVERPKPVVDGMAPMGWBGCS 540
Qy      541 RTGGGVGFSHRECKDPEPONGRCLGRRAKYOSCHTEBCPPDGSKFREOCCEKXNAYN 600
Db      541 RTGGGVGFSHRECKDPEPONGRCLGRRAKYOSCHTEBCPPDGSKFREOCCEKXNAYN 600
Qy      601 YTDMDGNLLQWVPKXAGVSPDRCLFCRARSEPFKFEAKVIDGTLGCEPTLAI CVRG 660
Db      601 YTDMDGNLLQWVPKXAGVSPDRCLFCRARSEPFKFEAKVIDGTLGCEPTLAI CVRG 660
Qy      661 QCVKAGCDHVVDSPKLDKCGVCGGKNSCRVSGSLPTVNGYNDIYTIIPAGATNIDVK 720
Db      661 QCVKAGCDHVVDSPKLDKCGVCGGKNSCRVSGSLPTVNGYNDIYTIIPAGATNIDVK 720
Qy      721 QSHSGVONDGNVYLLAKTADQYLLNGNLASAI BODILVKTILKYSGSIATLERLOS 780
Db      721 QSHSGVONDGNVYLLAKTADQYLLNGNLASAI BODILVKTILKYSGSIATLERLOS 780
Qy      781 RPLPEPLTVQLLTIVGEVFPKVKYTFVFPNDVDFSMOSKERRATNIIQPLLAHQWVLG 840
Db      781 RPLPEPLTVQLLTIVGEVFPKVKYTFVFPNDVDFSMOSKERRATNIIQPLLAHQWVLG 840
Qy      841 DMSBGSSTCGAGMORRTVECRDPGQASATCNKALKPEDAKPCESQLCPL 890
Db      841 DMSBGSSTCGAGMORRTVECRDPGQASATCNKALKPEDAKPCESQLCPL 890

```

## RESULT 2

```

US-09-989-687-4
; Sequence 4, Application US/09989687
; Publication No. US20040002449A1
; GENERAL INFORMATION:
; APPLICANT: Hassting, Gregg A.
; TITLE OF INVENTION: Meth1 and Meth2 Polynucleotides and Polypeptides
; FILE REFERENCE: 1488.107000D
; CURRENT APPLICATION NUMBER: US/09/989,687
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 890
; ORGANISM: Homo sapiens
US-09-989-687-4

```

```

Query Match      100.0%; Score 890; DB 11; Length 890;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 890; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MFPAPAPRMLPFLILLILLILLPLARGA PAPPAGQASL VVPTRLPGSAGELALHLSA 60
Db      1 MFPAPAPRMLPFLILLILLILLPLARGA PAPPAGQASL VVPTRLPGSAGELALHLSA 60
Qy      61 FGKGYLRLAPDSDSLAPFKIEBLGGSGRATGGRGIRGCFSGTUNGGEPSLSAAVSLC 120
Db      61 FGKGYLRLAPDSDSLAPFKIEBLGGSGRATGGRGIRGCFSGTUNGGEPSLSAAVSLC 120
Qy      121 RGLSGSFLDGEFFTIOPQAGAGSLAOPHRLQRMWGPAGARELPGRPEMEVETGEGORER 180
Db      121 RGLSGSFLDGEFFTIOPQAGAGSLAOPHRLQRMWGPAGARELPGRPEMEVETGEGORER 180
Qy      181 GDHODESEESQEEBAEGASEPPPLGATSRTRKFEVSEARFVETLLVADASMAAFYADL 240
Db      181 GDHODESEESQEEBAEGASEPPPLGATSRTRKFEVSEARFVETLLVADASMAAFYADL 240
Qy      241 QNHILITMSVAARIYKHPISIKNSINLMVYKYLIVDEKMGPEVSDNGGLTLRNFQNMQR 300
Db      241 QNHILITMSVAARIYKHPISIKNSINLMVYKYLIVDEKMGPEVSDNGGLTLRNFQNMQR 300
Qy      301 FNQPSDRPHBYDITALLTRQNFQCGEGICDTLGVADIGTICDPNKSQSVIIEDEGLQAAH 360
Db      301 FNQPSDRPHBYDITALLTRQNFQCGEGICDTLGVADIGTICDPNKSQSVIIEDEGLQAAH 360
Qy      361 TLAEHLGHVLSMPHDSKPCRTLFGPMGKHVMAPLFVHLNQTLLPMSPCSAWYLTLLDG 420
Db      361 TLAEHLGHVLSMPHDSKPCRTLFGPMGKHVMAPLFVHLNQTLLPMSPCSAWYLTLLDG 420
Qy      421 GHGDCILDPAGALPLPTGLPGRMALYOLDQCCROI FGPDRHCNPTSAQDVCAQLMCHT 480
Db      421 GHGDCILDPAGALPLPTGLPGRMALYOLDQCCROI FGPDRHCNPTSAQDVCAQLMCHT 480
Qy      481 DGAEPICHTKNGSLPMAOGTFCGPHLCSGSCLPPEEVERPKPVVDGMAPMGWBGCS 540
Db      481 DGAEPICHTKNGSLPMAOGTFCGPHLCSGSCLPPEEVERPKPVVDGMAPMGWBGCS 540
Qy      541 RTGGGVGFSHRECKDPEPONGRCLGRRAKYOSCHTEBCPPDGSKFREOCCEKXNAYN 600
Db      541 RTGGGVGFSHRECKDPEPONGRCLGRRAKYOSCHTEBCPPDGSKFREOCCEKXNAYN 600
Qy      601 YTDMDGNLLQWVPKXAGVSPDRCLFCRARSEPFKFEAKVIDGTLGCEPTLAI CVRG 660
Db      601 YTDMDGNLLQWVPKXAGVSPDRCLFCRARSEPFKFEAKVIDGTLGCEPTLAI CVRG 660
Qy      661 QCVKAGCDHVVDSPKLDKCGVCGGKNSCRVSGSLPTVNGYNDIYTIIPAGATNIDVK 720
Db      661 QCVKAGCDHVVDSPKLDKCGVCGGKNSCRVSGSLPTVNGYNDIYTIIPAGATNIDVK 720
Qy      721 QSHSGVONDGNVYLLAKTADQYLLNGNLASAI BODILVKTILKYSGSIATLERLOS 780
Db      721 QSHSGVONDGNVYLLAKTADQYLLNGNLASAI BODILVKTILKYSGSIATLERLOS 780
Qy      781 RPLPEPLTVQLLTIVGEVFPKVKYTFVFPNDVDFSMOSKERRATNIIQPLLAHQWVLG 840
Db      781 RPLPEPLTVQLLTIVGEVFPKVKYTFVFPNDVDFSMOSKERRATNIIQPLLAHQWVLG 840
Qy      841 DMSBGSSTCGAGMORRTVECRDPGQASATCNKALKPEDAKPCESQLCPL 890
Db      841 DMSBGSSTCGAGMORRTVECRDPGQASATCNKALKPEDAKPCESQLCPL 890

```

## RESULT 3

```

US-10-425-114-39107
; Sequence 39107, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabacka, Jack B
; APPLICANT: Cao, Yongwei

```

;; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
;; FILE REFERENCE: 38-21(53313) B  
;; CURRENT APPLICATION NUMBER: US/10/425,114  
;; CURRENT FILING DATE: 2003-04-28  
;; NUMBER OF SEQ ID NOS: 73128  
;; SEQ ID NO 39107  
;; LENGTH: 924  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATUERS:  
;; OTHER INFORMATION: Clone ID: LIB4654-025-F6\_F11 pep  
US-10-425-114-39107

Query Match 79.8%; Score 710; DB 12; Length 924;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 810; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 80 FKIERLGGSGRATGGERGLRGCFPSGTNGEPESLAVALSICRLGSGFLDGEFTIQPQ 139  
DB 114 FKIERLGGSGRATGGERGLRGCFPSGTNGEPESLAVALSICRLGSGFLDGEFTIQPQ 173  
QY 140 GAGGSIAQPHRLQRMWPGARPLPRGPEWEVETGEGOROBGDHDSSEESGEBAEGA 199  
DB 174 GAGGSIAQPHRLQRMWPGARPLPRGPEWEVETGEGOROBGDHDSSEESGEBAEGA 233  
QY 200 SEPPPLGATSRTRKFVSEARFVETLLVADASMAAFVADLQNHILLMSVARIYKHS 259  
DB 234 SEPPPLGATSRTRKFVSEARFVETLLVADASMAAFVADLQNHILLMSVARIYKHS 293  
QY 260 IKKSNIMVKKVLIVBEKKGPEVSDNGSLTLNFCQWQRFPOBDRPREHTDAILLT 319  
DB 294 IKKSNIMVKKVLIVBEKKGPEVSDNGSLTLNFCQWQRFPOBDRPREHTDAILLT 353  
QY 320 RQNFCCGEGLCDTLGVADIGTICDPNKSQVIEDGLQAAHTLAHLGVLSPHDSKP 379  
DB 354 RQNFCCGEGLCDTLGVADIGTICDPNKSQVIEDGLQAAHTLAHLGVLSPHDSKP 413  
QY 380 CTRLFGMGKHVMAPIFVHLNOTLPMSPCSAMITBELDGHGDCLLDAPGALPLPTG 439  
DB 414 CTRLFGMGKHVMAPIFVHLNOTLPMSPCSAMITBELDGHGDCLLDAPGALPLPTG 473  
QY 440 LPRGMALYQDQOCROQIFGPFRCPCNTSAQDVCAQJMCITDGAEPICHTKNSLPMADG 499  
DB 474 LPRGMALYQDQOCROQIFGPFRCPCNTSAQDVCAQJMCITDGAEPICHTKNSLPMADG 533  
QY 500 TPCGPGHLCSEGSCLPEEVEVERKPVVDGMAWPMGEGSRTCGGQVQFSHRECKDPP 559  
DB 534 TPCGPGHLCSEGSCLPEEVEVERKPVVDGMAWPMGEGSRTCGGQVQFSHRECKDPP 593  
QY 560 QNGRYCLGRARAYQSCHEECPDQKSPREOCCEKXNANYMDMGNLLQWPKKAYGS 619  
DB 594 QNGRYCLGRARAYQSCHEECPDQKSPREOCCEKXNANYMDMGNLLQWPKKAYGS 653  
QY 620 PRDRCKLFCRARSGSEKVEAVKIDGLTGPETLAIQVGGQCVKAGCDHVVDSPKLDK 679  
DB 654 PRDRCKLFCRARSGSEKVEAVKIDGLTGPETLAIQVGGQCVKAGCDHVVDSPKLDK 713  
QY 680 CGVCGGAGNSCRKXSGSLPTNGYNDIVITIPAGANTIDVKQSHGVQNDGYLALKTA 739  
DB 714 CGVCGGAGNSCRKXSGSLPTNGYNDIVITIPAGANTIDVKQSHGVQNDGYLALKTA 773  
QY 740 DGOYLNGNLAIASIEDILVKGITLKYSIATLEBLSFRLPEPLTVQLLTVEGEV 799  
DB 774 DGOYLNGNLAIASIEDILVKGITLKYSIATLEBLSFRLPEPLTVQLLTVEGEV 833  
QY 800 PKKYKTFEVPNDVDFMOSSEKERTATNIIQPLIHAQWVIGDMSSEGSTGAGMQRRTVE 859  
DB 834 PKKYKTFEVPNDVDFMOSSEKERTATNIIQPLIHAQWVIGDMSSEGSTGAGMQRRTVE 893  
QY 860 CRDPSGQASATCNALKPEDAKPCESQICPL 890  
DB 894 CRDPSGQASATCNALKPEDAKPCESQICPL 924

RESULT 4  
US-09-764-903-57  
Sequence 57, Application US/09764903  
Patent No. US2002030674A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: P7228  
CURRENT APPLICATION NUMBER: US/09/764,903  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PAM or file wrapper  
NUMBER OF SEQ ID NOS: 67  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 57  
LENGTH: 364  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-903-57

Query Match 29.6%; Score 263; DB 9; Length 364;  
Best Local Similarity 100.0%; Pred. No. 9,9e-232; Indels 0; Gaps 0;  
Matches 263; Conservative 0; Mismatches 0;

QY 530 WAPWPMGECGRTCGGQVQFSHRECKDPEPQNGRYCLGRARAYQSCHEECPDQKSPR 589  
DB 4 WAPWPMGECGRTCGGQVQFSHRECKDPEPQNGRYCLGRARAYQSCHEECPDQKSPR 63  
QY 590 EOCCKXNANYTMDGNLLQWPKKAYSPRDSCKLFCRARSGSEKVEAVKIDGLTLC 649  
DB 64 EOCCKXNANYTMDGNLLQWPKKAYSPRDSCKLFCRARSGSEKVEAVKIDGLTLC 123  
QY 650 GPETLAIQVGGQCVKAGCDHVVDSPKLDKCGVCGKNSCRKXSGSLPTNGYNDIVT 709  
DB 124 GPETLAIQVGGQCVKAGCDHVVDSPKLDKCGVCGKNSCRKXSGSLPTNGYNDIVT 183  
QY 710 IPAGANTIDVKQSHGVQNDGYLALKTAQGYLLNGNLAIASIEDILVKGITLKYSG 769  
DB 184 IPAGANTIDVKQSHGVQNDGYLALKTAQGYLLNGNLAIASIEDILVKGITLKYSG 243  
QY 770 SIATLERLQSFRLPEPLTVQLL 792  
DB 244 SIATLERLQSFRLPEPLTVQLL 266

RESULT 5  
US-09-918-171A-11  
Sequence 11, Application US/09918171A  
Patent No. US20020110894A1  
GENERAL INFORMATION:  
APPLICANT: Apte, Suneel  
APPLICANT: Hurekainen, Tiina L.  
APPLICANT: Hironaka, Satoshi  
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases  
FILE REFERENCE: 26473/04193  
CURRENT APPLICATION NUMBER: US/09/918,171A  
CURRENT FILING DATE: 2001-07-30  
Prior application data removed - consult PAM or file wrapper  
NUMBER OF SEQ ID NOS: 31  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 11  
LENGTH: 245  
TYPE: PRT  
ORGANISM: Homo sapiens ADAMTS-8  
US-09-918-171A-11

Query Match 24.4%; Score 217; DB 9; Length 245;  
Best Local Similarity 100.0%; Pred. No. 8.5e-190;  
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 196 AEGASEPPPLGATSRTRKFVSEARFVETLLVADASMAAFVADLQNHILLMSVARIY 255

```

DB 2 AEGAEPPPLGATSRTRFEVSEARFVETLLVADSMAPFAGADIQNHILTKSVARLY 61
QY 256 KHPISIKSNILNVAVVLLVEDEKMGPEVSDNGGLTLRNFQWQRFNQPSPDRHPEHYDTA 315
DB 62 KHPISIKSNILNVAVVLLVEDEKMGPEVSDNGGLTLRNFQWQRFNQPSPDRHPEHYDTA 121
QY 316 ILITRQNFQGGQGLCDTIGVADIGTICDPNKSCTVIEDEGLOAATTLAHELGHTLSMFD 375
DB 122 ILITRQNFQGGQGLCDTIGVADIGTICDPNKSCTVIEDEGLOAATTLAHELGHTLSMFD 181
QY 376 DSKPCTRLFGPMGKHVMAPLFVHNLQTLPMSPCSAM 412
DB 182 DSKPCTRLFGPMGKHVMAPLFVHNLQTLPMSPCSAM 218

RESULT 6
US-09-918-171A-9
; Sequence 9, Application US/09918171A
; Patent No.: US20020110894A1
; GENERAL INFORMATION:
; APPLICANT: Apce, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/04193
; CURRENT APPLICATION NUMBER: US/09/918,171A
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/369,364
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 9
; LENGTH: 905
; TYPE: PRT
; ORGANISM: Mus musculus ADAMTS-8
US-09-918-171A-9

Query Match
Best Local Similarity 3.9%; Score 35; DB 9; Length 905;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 617 GVSPPDRCKLFCRARGSEFFKVFPAKVDTGTCGP 651
DB 632 GVSPPDRCKLFCRARGSEFFKVFPAKVDTGTCGP 666

RESULT 7
US-09-802-582-8
; Sequence 8, Application US/09802582
; Publication No. US20020086354A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: 07334-323001
; CURRENT APPLICATION NUMBER: US/09/802,582
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 09/128,709
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: US 60/054,645
; PRIOR FILING DATE: 1997-08-04
; PRIOR APPLICATION NUMBER: US 09/130,491
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: US 60/054,966
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 09/388,280
; PRIOR FILING DATE: 1999-09-01

```

```

; PRIOR APPLICATION NUMBER: US 09/388,279
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-802-582-8

Query Match
Best Local Similarity 3.4%; Score 30; DB 12; Length 481;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 702 YGYNDIVTIPAGATNIDVQRSHPGVQNDG 731
DB 293 YGYNDIVTIPAGATNIDVQRSHPGVQNDG 322

RESULT 8
US-10-105-929-8
; Sequence 8, Application US/10105929
; Publication No. US20020137142A1
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/10/105,929
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/130,491
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-08-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/054,961
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Rattus rattus
US-10-105-929-8

Query Match
Best Local Similarity 3.4%; Score 30; DB 13; Length 481;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 702 YGYNDIVTIPAGATNIDVQRSHPGVQNDG 731
DB 293 YGYNDIVTIPAGATNIDVQRSHPGVQNDG 322

RESULT 9
US-10-365-227-8
; Sequence 8, Application US/10365227
; Publication No. US20030143632A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: 07334-323001
; CURRENT APPLICATION NUMBER: US/10/365,227
; CURRENT FILING DATE: 2003-02-12
; PRIOR APPLICATION NUMBER: US/09/802,582
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: US 09/128,709
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: US 60/054,645
; PRIOR FILING DATE: 1997-08-04

```

PRIOR APPLICATION NUMBER: US 09/130,491  
PRIOR FILING DATE: 1998-08-06  
PRIOR APPLICATION NUMBER: US 60/054,966  
PRIOR FILING DATE: 1997-08-06  
PRIOR APPLICATION NUMBER: US 60/058,108  
PRIOR FILING DATE: 1997-09-05  
PRIOR APPLICATION NUMBER: US 09/388,280  
PRIOR FILING DATE: 1999-09-01  
PRIOR APPLICATION NUMBER: US 09/388,279  
PRIOR FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 8  
LENGTH: 481  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-10-365-227-8

Query Match 3.4%; Score 30; DB 14; Length 481;  
Best Local Similarity 100.0%; Pred. No. 2e-18;  
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 702 YGYNDIVTIPAGATNIVKQSRHPGVQNDG 731  
DB 293 YGYNDIVTIPAGATNIVKQSRHPGVQNDG 322

RESULT 10  
US-09-321-987B-2  
Sequence 2, Application US/09321987B  
Patent No. US20020102210A1  
GENERAL INFORMATION:  
APPLICANT: Blueloch, Robert H  
TITLE OF INVENTION: Agent and Method for Modulating Cell Migration  
FILE REFERENCE: 960296.95386  
CURRENT APPLICATION NUMBER: US/09/321.987B  
CURRENT FILING DATE: 1999-05-28  
PRIOR APPLICATION NUMBER: 60/087,170  
PRIOR FILING DATE: 1998-05-29  
PRIOR APPLICATION NUMBER: 60/129,023  
PRIOR FILING DATE: 1999-04-13  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: Patent In Ver. 2.1  
SEQ ID NO 2  
LENGTH: 2150  
TYPE: PRT  
ORGANISM: Caenorhabditis elegans  
US-09-321-987B-2

Query Match 1.5%; Score 13; DB 9; Length 2150;  
Best Local Similarity 100.0%; Pred. No. 0.028;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGVQ 548  
DB 597 WGECSRTCGGVQ 609

RESULT 11  
US-09-800-729-155  
Sequence 155, Application US/09800729  
Patent No. US20020068319A1  
GENERAL INFORMATION:  
APPLICANT: Ni et al.  
TITLE OF INVENTION: 32 Human secreted proteins  
FILE REFERENCE: P2044P1  
CURRENT APPLICATION NUMBER: US/09/800,729  
CURRENT FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: PCT/US00/26013  
PRIOR FILING DATE: 2000-09-22  
PRIOR APPLICATION NUMBER: 60/155,709  
PRIOR FILING DATE: 1999-09-24

NUMBER OF SEQ ID NOS: 217  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 155  
LENGTH: 2165  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-800-729-155

Query Match 1.5%; Score 13; DB 9; Length 2165;  
Best Local Similarity 100.0%; Pred. No. 0.028;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGVQ 548  
DB 612 WGECSRTCGGVQ 624

RESULT 12  
US-09-963-791-22  
Sequence 22, Application US/09963791  
Patent No. US20020120113A1  
GENERAL INFORMATION:  
APPLICANT: Donoho, Gregory  
APPLICANT: Turner, C. Alexander Jr.  
APPLICANT: Friedrich, Glenn  
APPLICANT: Scoville, John  
APPLICANT: Zambrowicz, Brian  
APPLICANT: Sands, Arthur T.  
TITLE OF INVENTION: No. US20020120113A1 Human Proteases and Polynucleotides Encodi  
FILE REFERENCE: LEX-0105-USA  
CURRENT APPLICATION NUMBER: US/09/963,791  
CURRENT FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: US 60/169,769  
PRIOR FILING DATE: 1999-12-09  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 22  
LENGTH: 438  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-963-791-22

Query Match 1.3%; Score 12; DB 9; Length 438;  
Best Local Similarity 100.0%; Pred. No. 0.054;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGV 547  
DB 416 WGECSRTCGGV 427

RESULT 13  
US-10-419-276-22  
Sequence 22, Application US/10419276  
Publication No. US20030225258A1  
GENERAL INFORMATION:  
APPLICANT: Donoho, Gregory  
APPLICANT: Turner, C. Alexander Jr.  
APPLICANT: Friedrich, Glenn  
APPLICANT: Scoville, John  
APPLICANT: Zambrowicz, Brian  
APPLICANT: Sands, Arthur T.  
TITLE OF INVENTION: Novel Human Proteases and Polynucleotides Encoding the Same  
FILE REFERENCE: LEX-0105-USA  
CURRENT APPLICATION NUMBER: US/10/419,276  
CURRENT FILING DATE: 2003-04-17  
PRIOR APPLICATION NUMBER: US/09/963,791  
PRIOR FILING DATE: 2000-12-08  
PRIOR APPLICATION NUMBER: US 60/169,769  
PRIOR FILING DATE: 1999-12-09  
NUMBER OF SEQ ID NOS: 25  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 22

LENGTH: 438  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-419-276-22

Query Match 1.3%; Score 12; DB 12; Length 438;  
Best Local Similarity 100.0%; Pred. No. 0.054;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 536 WGECSRTCGGV 547  
DB 416 WGECSRTCGGV 427

RESULT 14  
US-09-803-589-10  
Sequence 10, Application US/09803589  
Patent No. US20020112251A1  
GENERAL INFORMATION:  
APPLICANT: McCarthy, Sean A.  
APPLICANT: Goodheart, Andrew D.J.  
TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING  
TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER  
FILE REFERENCE: 07334-325001  
CURRENT APPLICATION NUMBER: US/09/803,589  
CURRENT FILING DATE: 2001-03-09  
PRIOR APPLICATION NUMBER: US 09/128,709  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: US 60/054,645  
PRIOR FILING DATE: 1997-08-04  
PRIOR APPLICATION NUMBER: US 09/130,491  
PRIOR FILING DATE: 1998-08-06  
PRIOR APPLICATION NUMBER: US 60/054,966  
PRIOR FILING DATE: 1997-08-06  
PRIOR APPLICATION NUMBER: US 60/058,108  
PRIOR FILING DATE: 1997-09-05  
PRIOR APPLICATION NUMBER: US 09/388,280  
PRIOR FILING DATE: 1999-09-01  
PRIOR APPLICATION NUMBER: US 09/388,279  
PRIOR FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 518  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-803-589-10

Query Match 1.3%; Score 12; DB 9; Length 518;  
Best Local Similarity 100.0%; Pred. No. 0.063;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 706 DIVTIPAGATNI 717  
DB 291 DIVTIPAGATNI 302

RESULT 15  
US-09-802-582-16  
Sequence 16, Application US/09802582  
Publication No. US20020086354A1  
GENERAL INFORMATION:  
APPLICANT: McCarthy, Sean A.  
APPLICANT: Holtzman, Douglas A.  
APPLICANT: Goodheart, Andrew D.J.  
TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING  
TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER  
FILE REFERENCE: 07334-323001  
CURRENT APPLICATION NUMBER: US/09/802,582  
CURRENT FILING DATE: 2001-03-08

PRIOR APPLICATION NUMBER: US 09/128,709  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: US 60/054,645  
PRIOR FILING DATE: 1997-08-04  
PRIOR APPLICATION NUMBER: US 09/130,491  
PRIOR FILING DATE: 1998-08-06  
PRIOR APPLICATION NUMBER: US 60/054,966  
PRIOR FILING DATE: 1997-08-06  
PRIOR APPLICATION NUMBER: US 60/058,108  
PRIOR FILING DATE: 1997-09-05  
PRIOR APPLICATION NUMBER: US 09/388,280  
PRIOR FILING DATE: 1999-09-01  
PRIOR APPLICATION NUMBER: US 09/388,279  
PRIOR FILING DATE: 1999-09-01  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 16  
LENGTH: 551  
TYPE: PRT  
ORGANISM: Mus musculus  
US-09-802-582-16

Query Match 1.3%; Score 12; DB 12; Length 551;  
Best Local Similarity 100.0%; Pred. No. 0.063;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 706 DIVTIPAGATNI 717  
DB 324 DIVTIPAGATNI 335

Search completed: May 7, 2004, 12:06:24  
Job time : 55 secs





```

Db      361 TVCDPSRSCSYIEDDGLQAAFTTAHLEIGHVFNPMHDDAKHCASLINGVTDGSHLMSLS 420
Qy      400 LNCGLPMSPCSAWVLTLLDGGHDCCLDAPGALPLPTGLPGRMALYOLDQOCROI 459
Db      421 LDHQPMSPCSAWVLTLLDGGHDCCLDAPGALPLPTGLPGRMALYOLDQOCROI 478
Qy      460 DFRACPTNSAODVCAQWMC-HTDGAELPLCHTKNGSLPMADGTPCGPHGLCSBGSCLPEER 518
Db      479 ESKICPD--AASCTTILMCTGSGLLVCQTKH--FPMADQTSCEBGMCVSGCVNKT 534
Qy      519 VERKPVYDGMWPMWGECSRTCCGGVOPSHCECDPEPONGRCGLGRKAKYOSCHT 578
Db      535 MKHFAIPVHSGMGPWGDSCRTCCGGVOPSHCECDPEPONGRCGLGRKAKYOSCHT 594
Qy      579 EECPP-DGKSFREDOCEKXNAYNTDMDGN--LLQWPKYAGVSPRDRCLFCRARGSE 635
Db      595 EDCDNNKCTFREGECEAHNEFSKASF-GNEPTVEMTPKAGVSPKRCCLTCEAKGIGY 653
Qy      636 FKVEPAKVIDGTLGPEPLAICVRGQCYKAGCDHVDPSPKLDKCGVCGKNSCRVSG 695
Db      654 FFLVOPKVVDTGTPSPDSTSVCGQCYKAGCDRIIDSKKFKDCKGCGANGSTCKXMSG 713
Qy      696 SLTPNRYNDIVITIPAGATNIDVKORSHPGVQNDGNVLAALKTADGGYLLNGMLAISALE 755
Db      714 IYTSRPGYHIVITIPAGATNIEVKHNRGSRNMSFLAIRADGYILLNGMFTLSTLE 773
Qy      756 QDILVKGITLKYSGSIATLERLQSFRLPEPLTVQLLTPGEVFPKVKYTFVPRNDVF 815
Db      774 QDLTKYGTALRYSGSSALERIRSFPLKEPLTIQVLMV-GHALRPKIKFTYFM----- 826
Qy      816 SMGSKERATNIIQPLHQAQWVLDGMECSSTCGAGMORRTVECRPSGQASATCNAL 875
Db      827 ----KKKTESFNALPTF--SEWVIEHMGESCKTCSGMRVVOCCRDINHPSACEKAY 880
Qy      876 KPEDAKPCESQLCP 889
Db      881 KPASTRPCADLPCEP 894

RESULT 2
T00355
Hypothetical protein KIA0688 - human
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00355
D:NA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:98403880; PMID:9734811
A:Accession: T00355
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-837 <1SH>
A:Cross-references: EMBL:AB014588; NID:G3327189; PIDN:BA31663.1; PID:G3327190
A:Experimental source: Brain
C:Genetics:
A:Gene: KIA0688
C:Superfamily: thrombospondin type 1 repeat homology
F:519-576/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match      38.0%; Score 1842.5; DB 2; Length 837;
Best local Similarity 45.0%; Pred. No. 2.9e-117;
Matches 375; Conservative 130; Mismatches 244; Indels 85; Gaps 24;

```

Qy 116 AVSLCRG-LSGFILDGEFTIQ-----PQAGGSLAOPHRLQRWGAPARPLDRGPEW 168

```

Db      141 SLHMDGALLVLYQYRGELHLQPLEGGTPMSAGPGG--HILRRKSPAGC---QGMVC 194
Qy      169 EYETGEGRQROERGDHQEDSEESGEAEAGSEPPPLGATS-----RTKRVSSEARFVET 224
Db      195 NVKA-----PLGSPSPRBRKARFASLSRFVET 222
Qy      225 LTVADASMAFYGADLQNHILITMSVAARIYKHPISIKNSILMAYVKULIYDEKMGPEVS 284
Db      223 LVVADDKMAAFHGAQILKTYLLTVVAAAARAKHPSIRNPVSLVYTRVILISGSEGPVG 282
Qy      285 DNGGLTLNFCNMORFRNPQSPRHEHYDTAILLRQNFCCQOEGCDTLGVADIGTICDP 344
Db      283 PSAQDTLSFCAMQGLNTPEDSDPDHDTAILLFRQLCG-VSTCDTLGMADVGTVD 341
Qy      345 NKSCSVIEDEGLQAAHTLAHELGHVLSMPEDSKPCTRLFGPMG-KHHVMAFLPVHLNQT 403
Db      342 ARSCAIVEDDGLQSAFTAAHELGHVFNMLHNSKPCISLNGPLSTRHVMAPVMAHVDPE 401
Qy      404 LPMSPCSAMVLTLLDGGHDCCLDAPGALPLPTGLPGRMALYOLDQOCROI 463
Db      402 EPMSPCSAFITDFLDNGYGHCLDXKLPAPLHLPVTFPGKD--YADHQCCLTFGPDGRH 459
Qy      464 CENTSADVCAQWMC-HTDGAELPLCHTKNGSLPMADGTPCGPHGLCSBGSCLPEEYER 521
Db      460 CPQLPPP--CALMWSGHLNG-HAMQTKHS--PMADGTPCGPAQACMGGRCLHMDQLD 514
Qy      522 KPVVDGMAWPMWGECSRTCCGGVOPSHRECDPEPONGRCGLGRKAKYOSCHTEEC 581
Db      515 FNIPQAGGMPWGDSCRTCCGGVOPSHRECDPEPONGRCGLGRKAKYOSCHTEEC 574
Qy      582 PP-DGKSFREDOCEKXNAYNTDMDGN--LLQWPKYAGVSPRDRCLFCRARGSEPK 637
Db      575 PMSALTFREBQCAVYN--HRTDLKSPFGPMVPRATGVAPODQCLTCARALGYYY 632
Qy      638 VPEAKVIDGTLGPEPLAICVRGQCYKAGCDHVDPSPKLDKCGVCGKNSCRVSGSL 697
Db      633 VLEPRVVDGTPSPDSTSVCGQCYKAGCDRIIDSKKFKDCKGCGANGSTCKXMSG 692
Qy      698 TPTNRYNDIVITIPAGATNIDVKORSHPGVQNDGNVLAALKTADGGYLLNGMLAISALE 757
Db      693 RKFRVYNNVITIPAGATNIEVKHNRGSRNMSFLAIRADGYILLNGMFTLSTLE 750
Qy      758 ILVKGITLKYSGSIATLERLQSFRLPEPLTVQLLTPGEVFPKVKYTFVPRNDVF 810
Db      751 VVLPAGVSLRYSGATASSETLSGHGFLAQLPTLQVL-VAGNPQDTRLRYSPFVP 803

RESULT 3
T21371
Hypothetical protein F25H8.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T21371; T24896
R:Gajadaty, S.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19413
A:Accession: T21371
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <WIL>
A:Cross-references: EMBL:Z69360; PIDN:CAA93287.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: Clone F25H8
R:Gajadaty, S.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19949
A:Accession: T24896
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <WIL>
A:Cross-references: EMBL:Z69361; PIDN:CAA93288.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: Clone T13H10
C:Genetics:

```

A:Gene: CESP.F25H8.3

A:Map position: 4  
A:Introns: 31/1, 52/1, 135/2, 193/3, 216/1, 266/1, 495/2, 547/3, 564/3, 634/2, 744/1, 81

Query Match 30.4%; Score 1476; DB 2; Length 2165;

Best Local Similarity 34.1%; Pred. No. 7.4e-92;

Matches 327; Conservative 150; Mismatches 319; Indels 164; Gaps 29;

46 RLPSGAGELALHSAREKGFVRLAPDPSFL-----APEFIEHLGGS 88  
 96 RLQGVARD-----GHACHLRSDDAVYVHLHRMNOIPSHKXSVPHNSNFAM 148  
 89 GRATGGERLGR-----CFPSGTVNGEPESLAAVSLC--RLSGSFLLDGEETIQ 137  
 149 VLVDSEEVRRGMSRDPDCIYRAHVGVGHQ-STVNLCDSEDLGYGMALPFGHITVE 207  
 138 P-----QAGGSLAQPHRLQWGPAGARPLRGEWEVEIGEGQRORGRDH-QGDSBE 189  
 208 PISGNGTEHDGASRRHQHLVRKFDPMHFKSF-----DHLSNSTVN 248  
 190 ESOEEAEAGSEPPPLGATSRTRKRV-SEARFVETLLVADASMAAFYGADIQNHILTM 248  
 249 ETEETVATWQDQWEDVIERKASBRANSMDHYEVLVADTKGVEHGRSLADYVLTFF 308  
 249 SVAARIYKPSIKNSIMLAVYKVLVEDEKWEPEVSDNGGLTLRNCMQRFRNPQSDH 308  
 309 STVASIRHOSLRASINVVVVKLIVLTENAGPRITQNAQOTLDQFCRMQOYVNDPDS 368  
 309 PEHNDTILTLTRONFGQEGELCDPLGVAIGTICDPKSCSVIEDGLQAHATLAHELH 368  
 369 VOHNDVALILTRKDIGSGQCKDTLGLAELGTWCDMOKSCAILEDGLSAFTIAHELH 428  
 369 VLSMHPDSSKPCRLFGPMGK-----HHVAPLFLVHLNQTLPMSPCSA 411  
 429 VFSIFPHDDEKCS-TYMPVNVKCFQSGTKEDKTOFQNNFHIMAPLTLEYNHPWMSPCSA 487  
 412 MYLTLLDGGHG--DCLLDAPGAL--PLPTGLPERMALYQLDQCCROIFPPDRHCP 465  
 488 GMLRFLENRRGQTOCLFDQVEBRYEDVFAVEDEPKK--YAHQCKVFPPASELCP 545  
 466 NTSADVCAQALMCHT-DGAEPLCHTKNGSLPMADGTPCGPH--LCSGSC--DEEVE 519  
 546 ---YMPCRRLMCAFTFGSGMGRTQH--MPMADGTCDBSRSMFCHGACVLALESYL 600  
 520 ERKPVVDGKAPRPMGEGSRTGCGGVQSHRCKDPEPONGRYLGRARATQCHTE 579  
 601 K-----IDGQWGDWRSWGBCSRTOGGVQKLRDCSPKPRNGKTCVGGREERYSCNTQ 655  
 580 ECPEDGKSPFEOQCEKTN--AYNYTDMGNILQWPKXAGVSPDRCKLFCRARGRSEFK 637  
 656 ECPHDTPRYREVQSEFNKNDIGIGVASTNTHMVPKXANVAANERCKLCRLSGSAFY 715  
 638 VFEAKVIDTLCGETTALCYRGQCVAGCDHVVDSPRKLDKCGVCGKANSCKRYSGSL 657  
 716 LLRDKYVDGTPCDNRNGDICIYAGACMPAGCDHQLHSLTRDKCGVCGGDDSSCKVYKTF 775  
 698 TPT-NYGNDIVTIPAGATNIDVQRASHPGVONGNYALKTADGQYLNGNLAISAEQ 756  
 776 NEQGTFGNVEWKLIPASANIDIRQKGYNNKEDNYLSLPAANGERTLNGHFCVSLARQ 835  
 757 DILVKGTLIKYSGSIATLERLSFRPLPEPITVOLLTPGSEVPPKRYKTYF--VPNDV 813  
 836 QIARQDTLEYSGSDAIIERINGTPIRSIDYVAVLSVGH--PPDISYEMTAAVFNAY 893  
 814 -----DSM-OSSKERATYIIIOPL----- 832  
 894 IRPISALYLRVTDYTWECRACRGGQSQKLMCLDMSHTRQSHDRCCQVLRKQATRY 953  
 833 ---LHAQWVLGDMSEGSSTGAGMQRRTYECRDPG--QASAT--CNALKPEDAKPC 883  
 954 CNIDCSTWITEDEVSSCSAKCGSGQKQORVSCVMKEGDRQTPASEHLCDNRSKPSDIASC 1013

RESULT 4

T47158

hypoetical protein DKFZp762C1110.1 - human (fragment)

C:Species: Homo sapiens (man)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000C:Accession: T47158  
R:Blum, H.; Baurasachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.submitted to the Protein Sequence Database, March 2000  
A:Reference number: Z24379A:Accession: T47158  
A:Status: preliminaryA:Molecule type: mRNA  
A:Residues: 1-550 <AAA>A:Cross-references: EMBL:AL162080  
A:Experimental source: adult melanoma (Mawo cell line); clone DKFZp762C1110C:Genetics:  
A>Note: DKFZp762C1110.1Query Match 29.8%; Score 1444; DB 2; Length 550;  
Best Local Similarity 50.6%; Pred. No. 2.1e-90;

Matches 254; Conservative 92; Mismatches 134; Indels 22; Gaps 10;

391 HVMAPLFLVHLNQTLPMSPCSAMYLTLLDGGHGDCLDAPGALPLPTGLPERMALYQLD 450  
 11 HMMASMLNLSHQPSWSPCSAMITSLFDNNGHGLMDKQNPQLPGLDLP--TSYDAN 68  
 451 QCCROIPEPDPFHCPNTSADQVCAQLNC-HTDGAELPCLTKNGSLPMADGTPCGPHLCS 509  
 69 ROCQPTFEDBKCPD--AASTCTLMCTGTSGGLVQYQTH--FPMADGTSCEGKWC 124  
 510 EGSCLPEEVEVERPKVVDGKAPWPMGECSCRTGCGGVQSHRCKDPEPONGRYLCLGR 569  
 125 NKCYNKDKRKHFTLPFGSGMGMPGDCSRTGCGGVQYIMRBCDNPVPMGKXCGK 184  
 570 RAKYOSCHTECP--DGKSPFEOQCEKNAVNTDM--DGNLLQWPKXAGVSPDRCKLFC 627  
 185 RRYRSCNLEDCPDNNGKTFPEEOCEAHNEFSKSGFGSPAVEMI PKXAGVSPDRCKLI 244  
 628 CRARGRSEFKVFEAVIDTLCGETTALCYRGQCVKAGCDHVVDSPRKLDKCGVCGSG 687  
 245 COAKIGTFVLQPRVVDGTPCSPSTSVQGGCVKAGCRRITDSKKKFKCGVCGSG 304  
 688 NSCRVSGSLPTPNYNDIVTIPAGATNIDVQRASHPGVONGNYALKTADGQYLLNG 747  
 305 STCKKISGVSIAKGYHDITITPGATNIEVKQNRGSHNNGSLAIKADGTYLLNG 364  
 748 NLAIATIEDLIVKTLIKYSGSIATLERLSFRPLPEPITVOLLTPGSEVPPKRYKTYF 807  
 365 DYTLLSTLQDILMYKGVVLRYSGSSAALERIRSFPLKEPLTIOVLTV-GNALRPKIKTY 423  
 808 FVPNDVDSMOSKERATNTIIOPLHAQWVLGDMSEGSSTGAGMQRRTYECRDPGQA 867  
 424 FY-----KKKKEFPNI-PTFSA-WVLEWGECSKCELGWQRLVVECRDINGOP 471  
 866 SATCNKALKEPDAKPSQLCP 889  
 472 ABECAKEVPASTRCADHPCP 493

RESULT 5

T18517

procollagen N-endopeptidase (EC 3.4.24.14) I - bovine

N:Alternate names: procollagen N-proteinase  
C:Species: Bos primigenius taurus (cattle)C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T18517R:Collige, A.; Nielsen, B.V.; Lapiere, C.M.  
submitted to the EMBL Data Library, February 1996A:Description: Cloning of the cDNA of the bovine procollagen I N-proteinase.  
A:Accession: T18517A:Reference number: Z18941  
A:Status: preliminary;  
A:Molecule type: mRNA

translated from GB/EMBL/DDBJ

A:Residues: 1-1205 <COL>  
 A:Cross-references: EMBL:X56389; NID:e990769; PID:e228215; PIDN:CAA65253.1  
 A:Experimental source: skin  
 C:Genetics:  
 A:Gene: PC I-NP  
 C:Function:  
 A:Description: catalyzes cleavage of the propeptides of type I and II collagens prior to  
 C:Keywords: hydrolase; metalloproteinase

Query Match 22.7%; Score 1102; DB 2; Length 1205;  
 Best Local Similarity 30.9%; Pred. No. 9.6e-67;  
 Matches 316; Conservative 137; Mismatches 379; Indels 190; Gaps 41;

```

QY 3 PAPAAPRWL-PELLLLLLLLP---LARGA---PAPBAGGQASELVPTR-----46
DB 4 PAGAAGRLCPALLLLLLLPADARLAAAADPPGCGGAGARILAVPRDAQRLV 63
QY 47 -----LPG-----SAGELALHLSAFGKGFVRLAPDPSFL 76
DB 64 SHVSAATAPAGVTRRAAPAIPLSGSSEDDPGRIFFVTVFGDHLRLRNALV 123
QY 77 APEFKIERLGGSGRATGGERGLRGCFPGTVNGEPESTIA-AVSLCRGLSGSFLLDGEEPT 135
DB 124 ARGATVEMQESG-ATRVEPILGTLVGDVAGLASSVALSNODGLAGLRMEEEFF 182
QY 136 IOP--QGAAGSLAQPHLQ--RMGPAGARPLPRGEWEVETGEGORGERDQDESEES 191
DB 183 IEPLERKGLAAKAEQGRVHVYHRRPTSRPPLGGPQALDGG-----ISADSLD 233
QY 192 QE-----DEAGSAPPLPIGATSRKAFSEARF-VETLLVADASMAFYAGD-LONHI 244
DB 234 SRALGVLEBRVNSR-----RMKRHAADDYINIVLLGVDSDVQFQGTGHEVQKYL 285
QY 245 LTLMSVAARIYKHPISIKSINLMVVKLIVEDEKMGPEVS---DNGSLTLRNFQWQRR 300
DB 286 LTKMIVNEIYHDESLGHNINVLVIRILL---SYGKMSLIEIGNPQSLNCRWAYL 342
QY 301 FPGQDRDRPHVDTAILLTFQNGCGEGICDTLGVADIGTICDPKSGSVIEDGLQAH 360
DB 343 QOKPDTDEYHDAIFLTRDF--GPSGM--QGAIVTGMCHPFRSCTLHNEGFFSAF 398
QY 361 TLAEHGHVLSMHPDSDKPCRTLFGPMGKHVMARPLFVHLNQLTPMSPCSAMYLTELLDG 420
DB 399 VVAHETGHVLMGHEHGGQ--NRGDEVLGSLMARLVGAARFRFMWSCSGQESRXYL-- 454
QY 421 GHG-DCLLDAP---GAALPLPTGLPKMALYQLDQCRQIFGPRFRICPNTSAQDVCAQ 475
DB 455 -HSYDCLRDPTHTDMFALPOLPGLH-----YSNMEOCRFDGAGYMCATFRFTDPCKQ 508
QY 476 LMCHTDGAEPICHTKXGSLPMADGTPCGGHLCEGSGC--LPEEVEPRKPVVGGMAPW 533
DB 509 LMCSPHDNPFCKTKKG--PLIDGTMCAPGKHCPRGHCIWLPDLK-----DGNWAM 561
QY 534 GFWGCSRTGGGVOPSHRECKDPEPONGRYCLGRAKYQSCATEECPPDGKSFREQC 593
DB 562 SFFGCSRTCGTVFRTQCNDPANGARTCSGLAYDFOLCNSQDPCDYLADFREBOC 621
QY 594 EKXNAY-NYTDMDGULQWPKYAGVSPDRCKLFCRARGSEFEKFAKYLIDGLCG-P 651
DB 622 RQMDLYFEGDAQH--WLP--HEHRDAKERCHLTCESEKETEYVSMGRMHHDTGRCYK 677
QY 652 ETALICVRGQCVKAGCDHYVDSPRKLDKGVCGGKNSCRKYSG--SLTPTNYGNDIVT 709
DB 678 DAFSLCVRGDCRKGCDGIVIGSSKQEDKGVCGGNSHCKYKGFSSPKKLGLIKAFE 737
QY 710 IPAGATNIVKQSHNGVONDGNYLAKTAD--GQVLLANGNLASIAIEDDILYKGLIKYS 768
DB 738 IPAGASHLLIQE---ADTSHHLAVKMLKTKLLENDVDNSKTFIAMGVEMEYR 792
QY 769 GSIALTERLQSPRLPEPLTVOLLTVPGEVPPKVKYTFEVPND---VD-----814
DB 793 DEBGR-ETLQTMGFLGTLTV--LVIPEDARISLYKMYIHEDSLANDNNVLDDSVG 849

```

```

QY 815 -----FSMQSSKERATNIIQPLHLAQ 836
DB 850 YEMALKWSPCSCPGGSGQFTKYGCRRLRDLHKWVHVGCDQSVSKPKAIRTCNPOEQSG 909
QY 837 --WLGDMSECSSTGC-AGWQRRTYECDDPSQASAT-----CNKALKPEPAKCEQL 887
DB 910 PAVWVGEMPPCSCSGRTGMQVSVRCVQPLHNNTTTRSVHTKHGKND-A-RPGRRAKNREL 968
QY 888 CP 889
DB 969 CP 970

```

## RESULT 6

C89114  
 protein C37C3.6a [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 C:Accession: C89114  
 R:Anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio

A:Reference number: A75000; NUID:99069613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.banger.ac.uk/Projects/C\_e  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an  
 A:Accession: C89114  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1558 <STO>  
 A:Cross-references: GB:chr\_V; PIDN:AAC25867.1; PID:g3294501; GSPDB:GNO00023; CESP:C37C3.  
 A:Gene: C37C3.6a  
 A:Map position: 5

Query Match 11.6%; Score 562; DB 2; Length 1558;  
 Best Local Similarity 34.9%; Pred. No. 6.7e-30;  
 Matches 129; Conservative 51; Mismatches 136; Indels 54; Gaps 14;

```

QY 528 GGMWPMGPGSCRTGGGVOPSHRECKDPEPONGGRVCLGRAKYQSCATECPDGKS 587
DB 77 GNMGPVVPENECRSRSGGVQLKRCQSD-----CTGASVYISGNLMAC-ESGTD 127
QY 588 FREQCEKXNAYNYTMDGNLLQWPKYAGVSPDRCKLFCRARGSEFEKFAKYIDGT 647
DB 128 FRAEQCSKEND--EALDGNYHKMP-YKG--KNKCELVCPDESQNFYKADRVADT 180
QY 648 LCGPFTLACVAGQCVKAGCDHYVDSPRKLDKGVCGGKNSCRKYSGLTPTNY--GYN 705
DB 181 KCDKSNDCVQDEGCLPVQCDGKLSLFPDKCGKCTDQSGSTCKTIEGRFDRNLSPGH 240
QY 706 DIVTPAGATNIDVQKSHPGVONDGNYLAKTADGQVLLANGNLASIAIEDDILYKGL 765
DB 241 DILKPEGATNIIQE---ARKSTNNLALXNGSDHFLYNGN-GLIQEKEVEVGITF 294
QY 766 KYGSAITLERLOSFPPLPEPLTVOLLTVPGEVPPKVKYTFEVP--NVDMSMSKSR 823
DB 295 VYDD--AEPETLSAQPLSELTVALLFKRG-S-RDTALYFESIPLEBEVDY-----343
QY 824 ATNIIQPLHAQVWLGDMSECSSTCGAGMORRTVECRDPGQASA--TC--NKALKPE 878
DB 344 -----MYKDNMTPCSVSGKGVQTRNLVCLDGKQKGRVEDDLCEENATKPE 391
QY 879 DAKPCESQLC 888
DB 392 FEKSCETVDC 401

```

## RESULT 7

T34395  
 hypochlorite protein C37C3.6b - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 01-Dec-2000  
 C:Accession: T34395; T34394

R. Geisel, C., Bradshaw, H.  
 submitted to the EMBL Data Library, July 1996  
 A:Description: The sequence of C. elegans cosmid C37C3.  
 A:Reference number: 221518  
 A:Accession: T34395  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-2167 <GEN>  
 A:Cross-References: EMBL:U64857, PIDN:AAC25868.1, GSPDB:GN00023, CESP:C37C3.6b  
 A:Experimental source: Strain Bristol N2; clone C37C3  
 A:Accession: T34394  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1555 'SKF' <G52>  
 A:Cross-References: EMBL:U64857, PIDN:AAC25867.1, GSPDB:GN00023, CESP:C37C3.6a  
 A:Experimental source: Strain Bristol N2; clone C37C3  
 A:Gene: CESP:C37C3.6b; CESP:C37C3.6a  
 A:Map position: 5  
 A:Introns: 32/3; 104/2; 156/2; 207/1; 459/2; 536/3; 577/2; 1105/3; 1367/1; 1438/1; 1556/

Query Match 11.6%; Score 562; DB 2; Length 2167;  
 Best Local Similarity 34.9%; Pred. No. 1e-29;  
 Matches 129; Conservative 51; Mismatches 136; Indels 54; Gaps 14;

528 GGNAPGMPGBCSRTCGGVQFHSRECKDEPQNGRVCGRRAKYQSCHEECPPDGKS 587  
 DB GNMGPMPBENCSSCGGVQLEKQCSG-----CTGASVRIYISCHLNAIC-ESGID 127  
 588 FREQCCRYNAYNTMDGNLQWPKYAGVSPDRCKLFCRARGSEFFVFAKYIDGT 647  
 DB FRACQCKSFND---EALDGNVHKWTP--YKG---KMKCELCKPESGFFYYKMAKVVDTG 180  
 128 FRACQCKSFND---EALDGNVHKWTP--YKG---KMKCELCKPESGFFYYKMAKVVDTG 180  
 648 LCGSELLACVCGQCVAGCDHVVDSPKLDKCGVCGKGNCKRKYSGSLTPNMY--GVN 705  
 DB KCDKSNMDICVDGECPLVGCDCGKLGSSLPKPKCKCGDSCCTCTTGERDERNLSPGYH 240  
 706 DIVTIPAGATNIDVCKRSHPCQNDGNVYALKTADGQYLLNGNLASAIQDILVKGTL 765  
 DB DITLPREGATNIKIOE-----ARKSTNNLAKNGSDHFLYNGN-GLIQVEKEVEVGITP 294  
 241 DITLPREGATNIKIOE-----ARKSTNNLAKNGSDHFLYNGN-GLIQVEKEVEVGITP 294  
 766 KYSSSITLRLQGFRLPEPLTVQLLTVGGEVPPPKYKTFVFP--NDVDFSMQSKER 823  
 DB KYSSSITLRLQGFRLPEPLTVQLLTVGGEVPPPKYKTFVFP--NDVDFSMQSKER 823  
 295 VYDD--KEPFLSAQGPSEELTVALLFRKGS-RDTAKYFSTPLEEVDY----- 343  
 824 ATTNIQPLHAQVWLGPMSSECSSTCGAQWRTECRDPSGQASA---TC--NKALKPE 878  
 DB 344 -----MYKFNWTPCIVSCGKGVQTRRLYCIDGKNKGRVDDICENNAITKPE 391  
 879 DAKPCEGOLC 888  
 DB 392 FEKCEITVDC 401

RESULT 8  
 T18856  
 angiotensin inhibitor homolog - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
 C:Accession: T18856, T24653  
 R:McMurray, A.  
 submitted to the EMBL Data Library, July 1995  
 A:Reference number: Z19031  
 A:Accession: T18856  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1444 <W11>  
 A:Cross-References: EMBL:Z50004, PIDN:CAA90293.1, GSPDB:GN00028, CESP:CO2B4.1  
 A:Experimental source: clone CO2B4  
 R:McMurray, A.  
 submitted to the EMBL Data Library, July 1995  
 A:Reference number: Z19917  
 A:Accession: T24653

A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1444 <W12>  
 A:Cross-References: EMBL:Z50006, PIDN:CAA90302.1, GSPDB:GN00028, CESP:CO2B4.1  
 A:Experimental source: clone T07C5  
 C:Genetics:  
 A:Gene: CESP:CO2B4.1  
 A:Map position: X  
 A:Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 56

Query Match 10.9%; Score 529; DB 2; Length 1444;  
 Best Local Similarity 24.1%; Pred. No. 1.e-27;  
 Matches 236; Conservative 119; Mismatches 366; Indels 258; Gaps 48;

41 LVVPTRLGSGELHLHSARFGVFLAADDSTLAE-FKIELLGGSGRAT-GESEGL 98  
 DB 11 LEATRHPLKGNLTKMTFTAMNDTYHLNLKRSRIVSPHITSVAHGDDVTTAGLNDY 90  
 99 RCGPFGTVNGEPESIAAVSLCRGLSGSFLDGEFFIQ--POGAGSLSIAOPHRLQRMGP 156  
 DB 91 ECGHYQGEVYKSHGNKKAISDCGALMSIVMEDHFLVQLPK-----RVHHLQK--- 140  
 157 AGARPLPRGPEWETVGGQROEGDHOE---DSEESQEBEAGASEPPPLGATGRT 212  
 DB 141 -----EHLVYKRSAGLLTNABSKIRBEITRLQEBESFCDTSQDDPATTIPAHILF 194  
 213 KRFVS-----BARFY-----ETLVADASNAF-----YGADLQNH-----TLMSV 250  
 DB 195 NTITPISQALDSSTFPMMDDPTLEIGLFLDSKLFERHERYIQAEQHLEFSLALINN 254  
 251 AARIYKHSIKNSINLWVVKLYVEDEKMGPE-----VSDNGQL--TLRNFQWQRF 301  
 DB 255 VAVLYQOPLTFNLDIVLYRY-----EWMRTQPSALSTGVHKNQGAQSLIAFCRYQAHM 309  
 302 NQPSD-RHREHYDTAILTRQNFQGBGLCDTLGADIGTICDPKMSGVLEDEGLQAH 360  
 DB 310 NPGTLDTMNHDHVLVLLGYDI--YHTTTSVAGAPARAKCDBLPACSLVEGLHLSGF 367  
 361 TLAEHLGHVLSWPHD-----SKPCTRLFPGMKHNVAPLFVHLNQTLPMSFGSA--- 411  
 DB 368 VLAHEGMHGMVHDGVQNOCKGCCLSMAVAGAKT-----TMSDCSVREFN 415  
 412 MTLTELDGSHDCLLDA-PGAALPLPTG-----LPRMALYQDDCROIFGDFR-H 463  
 DB 416 AFLQLDSGRNCLRDASPGI--ISTNHSDLRLPQR--FLADQCCSTFWGADYXE 470  
 464 CPTGA-OPVCAQLMCHTDGAFLCHTRNGSLPMADGTPCGGHLCSGSCLPBEEVERP 522  
 DB 471 INGKAMDICRILWCGNSG-----TSTAHPLAEGSMCGANKKCHGQCTHMTFFGLTP 525  
 523 KPVVDGMAFWG-----PWGECSTRTCGGVQFHSRECKDEPQNGRVCGLRRAKYQSC 576  
 DB 526 VP-IDGESEWGAKEGCEPIQCAVSGSIYVGQHRDVCNPAFNNGGTTC----- 574  
 577 HTEECPPDGKSPREOC--EKTNAVYTDMD-GNLLQWPKYAGVSPDR----- 623  
 DB 575 -----EGANIRGLVCGATSSNCIGFTREBRGNKICSSIKXDPKPPQQLLGEFEST 627  
 624 --CKLFCRARGSEKVEAKYIDTGLCGPFTLACVGGQCYKACDHV-----DSFR 675  
 DB 628 QCRVWCHLIG-SELIRNGQFPDPTPGFID--ACVGGQCCIALSCDMKALVEQEDPR 684  
 676 -----KLDKGV--CGKGNCKRKYSGSLPTNNGVNDIVTIPAGATNIDVQK 721  
 DB 685 IGRSVHQMEESSMSSECSVSCGLGREVREKCS-----SKRKOQVSE 729  
 722 RSHP--GVQNDGNVYALKTADGQYLLNGNLASAIQDILVKGTLKTSGSA---TLER 776  
 DB 730 ESRPCSGVLRCDEEF-----GEW-----KEMGSCSEKCAAGV 761  
 777 LOSFRPLPEPLTVQLLTVGGEVFP-----KVXTTFVFP 810  
 DB 762 QKRFPC--LTDQCSKHLQEBRCPDNDCWTNMDWSSCGSGGGRRIYIRKCLDCK 818

```

QY 811 NDVDFMSQSKERATTTIIGPLHAQVNLGDMSCSTCGAGMORRTVECRDSSGQASAT 870
DB 819 CDGG---DLEESCNT---QXCISQSW--GDWLPSCVSGIGFOIRERLC---DGLCAT 867
QY 871 CNKALKEPAKPCESQOLCP 889
DB 868 ANK-----QARTCNQOQCP 881

RESULT 9
116892
hypothetical protein T19D2.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C/Accession: T16892
R/Bentley, D.
submitted to the EMBL Data Library, December 1995
A/Description: The sequence of C. elegans cosmid T19D2.
A/Reference number: Z18599
A/Accession: T16892
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-860 <BEN>
A/Cross-references: EMBL:U42846; NID:G1125809; PID:G1125810; PIDN:AA83600.1; CESP:T19D2
C/Genetics:
A/Gene: CESP:T19D2.1
A/Introns: 56/3; 96/1; 171/3; 199/3; 245/2; 293/2; 335/3; 395/3; 426/3; 484/2; 505/3; 58

Query Match 9.8%; Score 473.5; DB 2; Length 860;
Best Local Similarity 23.9%; Pred. No. 3.3e-24;
Matches 190; Conservative 109; Mismatches 284; Indels 213; Gaps 44;

QY 182 DH--QEDSEESQEEAAGASEPPPLGATSRTRFVSEAR-----FVETLLVAD 229
DB 90 DHEOCFDANDPYPDR-----ELAMKSMILRMVMDIRAPRRDILVELAVFAD 142
QY 230 ASM---AATGADLQNHILT-LMSVAARI---YKHSIKNSINLMVVKULYVD--EKV 279
DB 143 DAMMDHFKMYGKALENNMTITMAVNNIDVLTQRLLOPRINIKIVREILKNIPHLM 202
QY 280 GEVSDNGGL--TLNFCMQRFRNQPDRHEHYDTAIIITRONFCGSEGLCDTLGVAD 337
DB 203 NARKNSNDVRLDLAFQCYQNEINPRVADPRHMDHALFSQVDL--HRNGVKTVAGYAP 261
QY 338 IGTICDPKNSGVIDEGLQAAHTLAHELGHVLSNPHDSEKPCR---LFGP---MGKH 391
DB 262 VKGMSGVASCITINEGLDFGSVFVTHMGHSLGMYHDGNECDLRCCIMSPVSGSKTH 321
QY 392 VWAPLFLVHNLQTLPMSPCSAMYLTELDDGH--GD-----CLLDAPGALPL---PTGL 440
DB 322 -----WQCSVNEMATFV---GHGDFRFPNCLQDASNAEQGMVAFKSESP 365
QY 441 PGRMALYQLDQCCROI FGPDRH--CENTSADQVCAQJMC--HTDGAEPDLCTKNGLPWA 497
DB 366 PQQ--LFTLDEQCELFHGCWCHGHEKDGQIMQNICQWVWCNGSEGVIRTAH-----PAL 417
QY 498 DGTGPGPGHLCSESGCLPEEYERPKPVVDGMAWGMWGCSSRCG-----G 546
DB 418 EGTYGFGMTCRQGCQVSSQLMR---VTVGSMSTWN--DRPAPTCGRCGQCELRQIR 472
QY 547 VQFSRECKDPEFONGRCYCLGRARYQSCHTEECPDQ--KSFREOQCEK---NAVNYT 602
DB 473 IMRSIRQCNPSNNGARQCGDEARQVCHRDVNCGDSIERVATRVCSRLDEVALPNT 532
QY 603 DMDGNLLQWVPRYAGVSPDRCKLFCRARGSEFFKFAKVIDGLCGPELTALCVBQC 662
DB 533 IISGEMQF-----EQAMCKIWCLISGSTNITIV--SNFPDGAFCGG--QYIKGSC 581
QY 663 VKAGCDHVDPKDKGVCGGKNSCRKVGSLTPTNYGND-----IVTIP-- 711
DB 582 -----RPL-----LQGS-----TLAAVSEADCPPLVLTPTTPM 611

```

```

QY 712 -----AGATIDVKQRSHPGVQNDGNLALKTADGQYLNGNLALSAIBODILVKG 762
DB 612 PHVSVDPFAGKTNPYKHKTKTPELNEWSG-----SWASECVTDCHTQG 657
QY 763 TILKSSGISATL-----ERQSPRL---PEPLTVQLLVGGEVPPKVVYTFEPNDV 813
DB 658 VKVRVRCLAGVCAALRERQPTRPCTGSERPILT---SPQOTF--RRFIAPLR-- 710
QY 814 DFSQSSKERATTTIIGPLHAQVNLGDMSCSTCGAGMORRTVECRDSSGQASATCN 872
DB 711 -----RQTNMILRKVDH--W--GPWSACVTCGTGTLRRRENC---IGQBCA--- 751
QY 873 KALKPEDAKPCESQOLCP 888
DB 752 -----ETGFCVMOQC 761

RESULT 10
T00260
hypothetical protein KIAA0605 - human
C/Species: Homo sapiens (man)
C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C/Accession: T00260
R/Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.
DNA Res. 5, 31-39, 1998
A/Title: Prediction of the coding sequences of unidentified human genes. IX. The complete
A/Reference number: Z14086; M01D:98290545; PMID:9628581
A/Accession: T00260
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-951 <NAG>
A/Cross-references: EMBL:AB011177; NID:G3043733; PIDN:BA25531.1; PID:G3043734
A/Experimental source: brain
C/Genetics:
A/Gene: KIAA0605
C/Superfamily: thrombospondin type 1 repeat homology
F/46-106/Domain: thrombospondin type 1 repeat homology <THR1>

Query Match 9.0%; Score 438; DB 2; Length 951;
Best Local Similarity 31.7%; Pred. No. 9.8e-22;
Matches 113; Conservative 45; Mismatches 124; Indels 74; Gaps 14;

QY 526 VDGG-----WAPWPGECSSRTGCGGVQSFHRC-----KDPFONGRYTLGRARYQ 574
DB 39 LRGTDTATFWMGEMTKMTAFSRSCGGVTSQERHCLQRRKSVGPGARRCTGTSKRYQ 98
QY 575 SCITECPDQKSPFREQCEKYNAYNTDGNLLQWVPKYAGVSPRD-----RCKLF 627
DB 99 LCRVQCEPDPGRSFRBQCVSNHNY---NRTTHQMKFLY---PDDVHHSKPCDLH 151
QY 628 C-RARGSEPKVFEAKVIDGLCGEPLT-AICVRQCVYAGCDHYVDSFRKLDKGCVCQ 685
DB 152 CTTVDQGRQLMW---PARQTSCKLTLRGVCVSGKEPILGCDGVLFSTHLDKCGICQ 208
QY 686 KNSCKKNSGSLPTN--YGYNDIYTIIPAGANIDVQKSHGVQNDGYALAKTPADQY 743
DB 209 DGSSTHVTGNRKGAHLGYSLVTHIFAGARDIQVER-----KKSADVLALDEAGY 263
QY 744 LINGNLAISAILBODILVKGTILKYSGSI-----ATLERLOSFPPLPEPLTVQ----- 790
DB 264 FRNGNKKVDS--PNNFNIAGIVVYKRRPMDVETGRIYVIAQGFPTQGLVWVWVWNGKSP 322
QY 791 -----LTVPEGVFPKVKYTF-----FVNDVDFMSQSKER 823
DB 323 SITFEYTLLOPFHESRPPITYGFSSASQGLDAGLMGFLPHNSGLYGASSER 378

RESULT 11
T11576
hypothetical protein F08C6.1 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C/Accession: T11576

```



```

QY 565 YCLGRAPKQSCHECECPDGSFREQCEKXNAVNTYMDGNLLQWVPKYAGVSPDRDC 624
DB 472 LCRGR-----TSECD-----VPSE-----485
QY 625 KLFCAARGSEPKVFEAVIQTGLGPEPLAIQVRGQC--VKAGCDHVSPRKL-----677
DB 486 -----CNSSQPCQPDVFILQNGYPC-QNNKAYCTNMGQYVDAQCYVFGSKAPAPDC 539
QY 678 -----DKGVCCKGKNSCRK-----VSGSLTPFN-----YG-YNDIVTIPAGAT- 715
DB 540 FLEVNSKQDRFNGCFSNGEYKKCATGNALGKLCCEVNGEIPVFGIIPALIQPPSHGTG 599
QY 716 --NIVKQRS--HPGVQNDG 731
DB 600 CMGVDFOLGSDVDFDGMVNEG 620

```

## RESULT 13

monocyte surface antigen MS2 precursor - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 03-Feb-1993 #sequence\_revision 03-Feb-1993 #text\_change 26-Aug-1999  
 C/Accession: A60385  
 R.Yoshida, S.; Setoguchi, M.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.  
 Int. Immunol. 2, 585-591, 1990  
 A/Title: Molecular cloning of cDNA encoding MS2 antigen, a novel cell surface antigen  
 A/Reference number: A60385; PMID:91197896; PMID:1982220  
 A/Accession: A60385  
 A/Molecule type: mRNA  
 A/Residues: 1-826 <YOS>  
 A/Residues: 1-826 <YOS>  
 A/Cross-references: EMBL:X13335  
 C/Superfamily: mouse melanin alpha; disintegrin homology  
 C/Keywords: glycoprotein; surface antigen; transmembrane protein  
 F/1-14/Domain: signal sequence #status predicted <SIG>  
 F/402-484/Domain: disintegrin homology <DIS>  
 F/659-683/Domain: transmembrane #status predicted <TM>  
 F/330/Active site: Glu #status predicted

Query Match 6.4%; Score 309.5; DB 2; Length 826;  
 Best Local Similarity 22.4%; Pred. No. 4,5e-13;  
 Matches 171; Conservative 85; Mismatches 292; Indels 215; Gaps 36;

```

QY 14 LLLLLLLPLARAPAPAGQASLVPTPLPGSAGELAL-----HISA 60
DB 4 LMLSLVLTMPAV--APGPPLPHVKQYEVWPRRLAASRRALPSHWQYFESLYALGT 61
QY 61 FKGIVLPLAPDPSFLAPFKIERIGSGG-RATGGERGLRGCFPGTVNGEPESLAIVSL 119
DB 62 SHVFTLHKRKRDLLGSSYETTYSAANGSEYTELQDQDHLQGHVEGREGSAIST 121
QY 120 CGLSGSFLLDGEETIOP-----QAGGSLAOPHRLQWGPAGARPL-----PRGEW 168
DB 122 CAGLGAFRFVGVSTVHLIEPLDADEGQAMVQAKHLQKAGTCVQKDTNINLDPRALE- 180
QY 169 EYETSGQRQERGDQDEDESESGEBEAGASEPPPLGATSRTRFVSEARFETLLVA 228
DB 181 -----YRAOPRMWLERETRYELVVA 204
QY 229 DASMAFYGA--DLQNHILTMSVARIYKPSIRNSINLMVVKLYIYEDKMGPE--V 283
DB 205 DQGFQKQSGRAVAVQRLVNVNVDKLYQELSF-----RVLVGLGLIMWKDKYI 255
QY 284 SNGGILTLNFCMQRFRQPSDRPHRYDTAILITRONFCQGESLCTTLGVAIDIGTICD 343
DB 256 SRYAVNTLENFLSMBEONLQ--GQHP--HDNVQILTGVDFIG-----STVGLAKVASLCS 306
QY 344 PKKSGCVID---EGLOAHTLAHELGHVLSVPHDSKP-CTRLFGPKGKHVAVAPLVH 399
DB 307 -RHSAGVNDHKSNGISVASTVAHLEGNLGSHEBDIPGC---YCPREPGGGGIMES 362
QY 400 INQTLIP--WSPCSAMYTELLDGGHGDCLLAPGALPLPTGLPERMALYOLDQ-----451

```

```

DB 363 IGSKEPRIRSCSKIDLESFVTKPQTGCLTNVPDYNRFYGVCGNLFYHGEQDCGTP 422
QY 452 -----OQRQIFGPD-----FRHCPTSAQDVC-----AGLWCHTGAEL 486
DB 423 QDCQNPCCNATTQQLYKGAEBGASGTCHECKVKPAAGEVCRLSKDCDLEEFQ--DGRKPT 480
QY 487 CHTKNSLP-----MADGTPCGRHLCBSGSCPESEVEPRKRVVDGMAPMGWGECSIT 542
DB 481 C-----PEDAFQNGTFC-PGYCFDPSCTTLAQCCDL-----WGP-----516
QY 543 CGGQVQFSHRECKDPEPQNGRGYCLGRARAKYQSCHECECP--DKSF--REQCEKXNA 598
DB 517 -----GARVAADSCVTFSPPCQNGRMVSGRLNRCAL--549
QY 599 YNTYMDGNLLQWVPKYAGVSPDR--CLT-----FCRARGR--SEFKVFEAVIQTLCG 650
DB 550 --YCE-----GGQKPLERSFTFSNNHGVCHALGTGNSITDFEL-VLQGTIC- 593
QY 651 PETLACVAGQCVKAGCDHVVDSPRLDKGCVGCKGN-SCRK 692
DB 594 -EBGKVCMDGSCQDRLRVYRSENGSAKNNHGVCHNKBCHCHK 635

```

## RESULT 14

hypothetical protein F53B6.2 - Caenorhabditis elegans  
 C/Species: Caenorhabditis elegans  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C/Accession: T22545  
 R.White, S.  
 submitted to the EMBL Data Library, October 1996  
 A/Reference number: Z19578  
 A/Accession: T22545  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 11059 <WLL>  
 A/Cross-references: EMBL:Z61086; PIDD:CAMB0121.1; GSDDB:GN00019; CESP:F53B6.2  
 A/Experimental source: clone F53B6  
 C/Genetic:  
 A/Gene: CESP:F53B6.2  
 A/Map position: 1  
 A/Intons: 38/3; 92/3; 131/3; 169/1; 236/2; 282/3; 349/2; 376/2; 420/1; 576/3; 759/3; 8

Query Match 6.4%; Score 309.5; DB 2; Length 1059;

Best Local Similarity 24.1%; Pred. No. 6.1e-13;  
 Matches 91; Conservative 45; Mismatches 113; Indels 129; Gaps 14;

```

QY 530 MAPGPMGSCSRTCGGQVQFSHRECKDPEPQNGRGYCLGRARAKYQSCHECECPDGSFPR 589
DB 26 WAMNSFMSGCTKTCGGGVSRQLRLCTSK-----CSGESVRFKCAQKTESRRLAR 78
QY 590 EQCEKXNAVNTYMDGNLLQWVPKYAG--VSPDRCKLPCPAR--GRSEFKFEAKYVD 645
DB 79 DTIC-----GGEIIVSRGQCEVVCSSRLTGAN-----FLWRVVD 112
QY 646 GTLCGPET-LAICVRGQCVKAGCDHVVDSPRLDKGCVGCKGNSCR-----KYSGSLT 698
DB 113 GTPCQATSRVAVSSKSCQIVGCDGLISSFFPDAGVCGGSGDPTDNDKFTMKYSEET 172
QY 699 PTNYGVNDIVTIPAGATNIDVQRSHPGVQNDGNTALATADQVLLNGLAISALEQDI 758
DB 173 -----ACASND-----179
QY 759 LVKGTILTKSG--SIATLERLOSFRLEPELTVQLLTVGSEVPEPKVATYTFVPRVDVF 815
DB 180 ---DIVDWSGASRSTASTS-----QPIVVCNATITKRVPEKL-----CADK 218
QY 816 SMOSSKERATTNIIQPLIHAQWTLGDMSECSSTCGAGMQRRTVECHDPSGQAS-----AT 870
DB 219 LRKVEARPCPMILIC--SRWMAADWTECVPHCGGTRKREYVCVQTANHTVHVDPTE 275
QY 871 CNKALKEPDAKPECSQIC 888

```



Db 276 CENGTAPAEENCSTSC 293

## RESULT 15

G02390  
disintegrin-like metalloproteinase MDC15 (BC 3.4.24.-) - human  
C/Species: Homo sapiens (man)  
C/Date: 21-Dec-1996 #sequence\_revansion 06-Jun-1997 #ext\_change 31-Dec-2000  
C/Accession: G02390; PC4263  
R/Herron, B.; Raines, E.W.; Ross, R.  
submitted to the EMBL Data Library, January 1996  
A/Reference number: H01157  
A/Accession: G02390  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA  
A/Residues: 1-814 <HER>  
A/Cross-references: EMBL:U66005; NID:g1335871; PIDD:AAC5112.1; PID:g1335872  
R/MKle, N.; Edwards, T.; Dallas, D.O.; Houghton, A.; Stringer, B.; Graham, R.; Russell, J.  
Biochem. Biophys. Res. Commun. 230, 335-339, 1997  
A/Title: Expression of members of a novel membrane linked metalloproteinase family (ADAM  
A/Reference number: PC4263; MID:97168971; PMID:9016778  
A/Accession: PC4263  
A/Status: nucleic acid sequence not shown  
A/Molecule type: mRNA  
A/Residues: 1-461 <MCK>  
A/Experimental source: articular chondrocyte  
C/Comment: This protein is a membrane bound protein and involved in cell/cell and cell/m  
C/Superfamily: mouse metrin alpha; disintegrin homology  
C/Keywords: hydrolase; metalloproteinase; zinc  
F/420-503/Domain: disintegrin homology <Dis>  
F/348,352,358/Binding site: zinc (His) #status predicted  
F/349/Active site: Glu #status predicted

## Query Match

6.3%; Score 306; DB 2; Length 814;

Best local Similarity 25.1%; Pred. No. 7.6e-13;  
Matches 181; Conservative 75; Mismatches 273; Indels 192; Gaps 42;

QY 42 VVPTRLPGSAGELALHLSAFKGFVTLAPDDSLAPFEKIERLGSGPAT-----GGE 95  
Db 60 VLQTSLP---EPRIKIKELDGDSDHILQLNRE-LVP-----GRFTLWYQPDGT 105  
QY 96 R-----GLRGCFPSGTNVNPEPSLAVSICRLSGSFLDGEFTIOPGAGSLAQP- 148  
Db 106 RVVSEGHITLNCYQGRVRYAGSWSICTCSGLRGLVUTPRSTYTLQGRDLOGPPI 165  
QY 149 -HRLQWGPAGAPLPGRPEWEVETGEGQERGDHODESEESQEEAEAGASEPPPLG 207  
Db 166 ISRIQDLHLPG-----HTCALSWRESVHTQTPPEHPLG 198  
QY 208 A--TSRTKRFVSARFVETTLVADASMAAFYGLDNLHILTMVSVAARI--YKHPISIKNS 263  
Db 199 QRIRRRRDVVTETKVELIVADHSEAKRY-RDQHLNRTLEVALLDLTFRR----- 252  
QY 264 INLMVAVLIVDEKNGP---EVSNGGLTLNFCNMORRFNQPSPDRPHNYDAIILT 319  
Db 253 ---LNVVALVGLAEAWTORDLVEISPNPAVTLENFLHWRRAHLRP--RLP--HDSAQLV 305  
QY 320 RQNPCCGEGLCDTLGVADITICDPNKSQSVIEDEG---LQAHTLAHELGHVLSMHPD- 375  
Db 306 GTSFSG---PTVGMAIIONISCPDPSGVMNDHSTISIGVASTAHELGHSLGLDHD 360  
QY 376 --DSKPTCLFLGPMGKHVVAPLFVHLNQLTFWSPCSAMYLTEILDGHDCLLDAPGAA 433  
Db 361 PGRNSCPSP---GAPAKTCIMEASTDPLPGLNFSNCSRRALBKALDGMGSCIFER---- 413  
QY 434 LPLPTGLPGMALYQLDQCRQIFGPDPRHCPTNSADVCAQLMCHTDGAEPLCHTKNGS 493  
Db 414 --LPSLPP--MAAF---CGNMVFEPEGQCD-----CGFL--DDCVDPCC----- 448  
QY 494 LPMWADGTPC--GPGHLC--SEGSCLPEEEVERPKPVVDGGAFWGPMGEGCS--RTCGGQVQ 548  
Db 449 ----DSLTCQLRPAQCAADGPGCCQNCQL-RP-----SGWQGRPTRGDDCLPFCPG--- 495

QY 549 FSHRECK-----DPEPONGGR-YCL-GRAPKY-QSCHTEBCP---PDGKSFREQQCEKY 596  
Db 496 -DSQCCEPDVSLGGEFCACAGQAVCMHGRCASTYQCCQSLMGPAQAPALPLCLQTANTRG 554  
QY 597 NAYNYTMDGNLQWVERKTAGVSPDR-C-KLFCRANGSE-----FKVFEAK-- 642  
Db 555 NAFSCGRNP8-----GSYVSCTPRDAICGQLQCOT-GRTOPLGSIKDLMEITDVNGT 608  
QY 643 -----VIDGTLGPETIAICVAGQCVKAGCDHVVDSPRLDKCGV 682  
Db 609 ELNCSWHLIDGSDVAPPLTLPTGACPG--LVCIDHRCQVDLGAQECRSKCHGAV 666  
QY 683 C 683  
Db 667 C 667

Search completed: May 7, 2004, 11:51.41  
Job time : 30 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: May 7, 2004, 11:46:40 ; Search time 51 Seconds

(without alignments)  
5506.102 Million cell updates/sec

Title: US-09-989-687-4

Perfect score: 4853

Sequence: 1 MFPAAPAPRWLPFLLLLL.....CNKALKPEDAKPCESQLCPL 890

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTRMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_prokaryote:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2320.5	47.8	967	4 Q8NE26	Q8NE26 homo sapien
2	2112.5	43.5	759	6 Q8H2M8	Q8H2M8 equus caball
3	1813	37.4	833	11 Q8K384	Q8K384 mus musculus
4	1813	37.4	845	11 Q8BNJ2	Q8BNJ2 mus musculus
5	1794.5	37.0	839	6 Q7E9S5	Q7E9S5 bos taurus
6	1483.5	30.6	623	11 Q8BGP4	Q8BGP4 mus musculus
7	1475	30.6	2165	5 Q19791	Q19791 caenorhabditis
8	1198	24.7	1688	5 Q8SXB0	Q8SXB0 drosophila
9	1182	24.4	1059	5 Q9W493	Q9W493 drosophila
10	1161.5	23.9	1009	11 Q8KX11	Q8KX11 mus musculus
11	1161.5	23.9	1600	11 Q8C1B3	Q8C1B3 mus musculus
12	1125.5	23.2	1070	11 Q8CG28	Q8CG28 mus musculus
13	1067.5	22.0	1092	11 Q8BKX1	Q8BKX1 mus musculus
14	990	20.1	900	11 Q8K206	Q8K206 mus musculus
15	976	20.1	769	5 Q8KRL5	Q8KRL5 drosophila
16	836	17.2	183	11 Q9JCP1	Q9JCP1 mus musculus

17	830	17.1	1427	4 Q9G137	Q9G137 homo sapien
18	740.5	15.3	269	6 Q9GL54	Q9GL54 oryctolagus
19	683.5	14.1	340	11 Q91256	Q91256 mus musculus
20	614.5	12.7	1280	11 Q9EPX2	Q9EPX2 mus musculus
21	592.5	12.2	203	11 Q8CH80	Q8CH80 cavia porce
22	579.5	11.9	1235	4 Q95428	Q95428 homo sapien
23	571	11.8	1572	5 Q44938	Q44938 haemophilus
24	562	11.6	1487	5 Q8MPV5	Q8MPV5 caenorhabditis
25	562	11.6	1558	5 Q81710	Q81710 caenorhabditis
26	562	11.6	2167	5 Q7E840	Q7E840 caenorhabditis
27	547	11.3	3198	5 Q9U8G8	Q9U8G8 manduca sex
28	541	11.1	192	6 Q95N24	Q95N24 equus caball
29	541	11.1	2174	5 Q9GQRO	Q9GQRO drosophila
30	541	11.1	2776	5 Q869A0	Q869A0 drosophila
31	541	11.1	2898	5 Q86829	Q86829 drosophila
32	536	11.0	2772	5 Q9W4V4	Q9W4V4 drosophila
33	529	10.9	1461	5 Q8V7A8	Q8V7A8 caenorhabditis
34	498	10.3	377	11 Q8BL10	Q8BL10 mus musculus
35	495.5	10.2	439	4 Q8N6G6	Q8N6G6 homo sapien
36	493.5	10.2	766	4 P82987	P82987 homo sapien
37	490.5	10.1	1091	5 Q9M1Z6	Q9M1Z6 drosophila
38	488.5	10.1	525	4 Q86RW4	Q86RW4 homo sapien
39	482	9.9	339	4 Q8NEK2	Q8NEK2 homo sapien
40	473.5	9.8	872	5 Q22580	Q22580 caenorhabditis
41	471.5	9.7	497	11 Q8B2D1	Q8B2D1 mus musculus
42	453.5	9.3	1036	11 Q80T21	Q80T21 mus musculus
43	450	9.3	187	6 Q86TH1	Q86TH1 homo sapien
44	446	9.2	446	6 Q95N23	Q95N23 equus caball
45	444.5	9.2	957	11 Q7TSK7	Q7TSK7 mus musculus

## ALIGNMENTS

### RESULT 1

Q8NE26 ID Q8NE26 PRELIMINARY; PRT; 967 AA.  
AC Q8NE26; 01-OCT-2002 (TRENBLREL. 22, Created)  
DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update)  
DT 01-OCT-2003 (TRENBLREL. 25, Last annotation update)  
DE Hypothetical protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OC NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strubeberg R.;  
RL Submitted (Aug-2002) to the EMBL/GenBank/DBJ databases.  
DR GO; GO:0004222; F:metallopeptidase activity; IEA.  
DR GO; GO:0008270; F:zinc ion binding; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR006586; ADAM\_cysteine.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR InterPro; IPR001590; Peptidase\_M12B.  
DR InterPro; IPR002870; Peptidase\_M12B\_N.  
DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR008085; TSP 1.  
DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
DR Pfam; PF01421; Repolysin; 1.  
DR Pfam; PF00907; tep\_1; 3.  
DR PRINTS; P011705; TSP1REPEAT.  
DR SMART; SMO0608; ACR; 1.  
DR SMART; SMO0209; TSP1; 3.  
DR PROSITE; PS0215; ADAM\_MEPPO; 1.  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
DR PROSITE; PS0092; TSP1; 3.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.

KW Hypothetical protein.  
SQ SEQUENCE 967 AA, 105387 MW, FFIID39674201C3D CRC64;

Query Match 47.8%; Score 2320.5; DB 4; Length 967;  
Best Local Similarity 49.6%; Pred. No. 6,7e-185;  
Matches 451; Conservative 146; Mismatches 250; Indels 63; Gaps 22;

```

QY 11 LPELILLILLPLLRGAAPAPAGGQASELVV--TLPSAGEMLHLISAFKGVLR 68
DB 33 VFTLLILAAALIAVS-DALGRPS--EDBELVPELEAPGH-GTTRRLHAFPGQDLE 88
QY 69 LAPDUSFLAPEFKIERLGSGSGRATG---ERLRCGFSGTYNGEPESIAVLCRGL 123
DB 89 LRPDSSFLAPGFTLQNV---GRKSGSTPLPETDLAHCFYSGTVNGDPSSAALSLCEGV 145
QY 124 SSGFLLDEEFTIQOGAGSLAOPHRLQRMGPAPARLP-----RGPEWEY---ET 172
DB 146 RCAFYLLGSAVFIQPLPAAS---ERLATAAPGKPPAPLQFHLLRRNRQDVGTCGV 200
QY 173 GEGQROERGDHQDESEESOEERAEAGASEPP-----PLGATS-RTKRFVSEARFV 222
DB 201 VDDEPRPTGKATEDEDGTEDEGEDEGPQMSPODPALQGVGPTGSGIRKKRFVSSHRYV 260
QY 223 ETLIVADDSMAAFYQADIQNHILITMSVAARIYKSPISKINIMVYKVLIVDEBKGPFE 282
DB 261 ETLIVADDSMAEFHSGKHYLLTLFVAARLYKHPISIRNSVSLVYKVLIVHDEKGPFE 320
QY 283 VSDNGGLTLRFNCNRRFPNCPDRHPHYDTAILLTRNFCQGEGLCDTLGVADIGTIC 342
DB 321 VTSNAILTRNFCNRRFPNCPDRHPHYDTAILLTRNFCQGEGLCDTLGVADIGTIC 379
QY 343 DPNSGCSVIEDEGLQAAHTLAHELGHVLSMPPHDSKPTCTRLFGPMGKHVYAPLFHLNQ 402
DB 380 DPNSGCSVIEDEGLQAAHTLAHELGHVLSMPPHDSKPTCTRLFGPMGKHVYAPLFHLNQ 439
QY 403 TLPMSGCSAMYTELTDGGHDPCLLDAPGALPLPTGLPGMYALYQLDQCCROQFGPDR 462
DB 440 SGPMSBSCAYMITSLFDNGHSECLMDKQNPQLQDGLDPG--TSDANRQOQFTFGEDSK 497
QY 463 HCPNTSAODVCAQLWC-HTGAEPICHTKNGSLPMADGTPCGPGLCSGSGCLDEEVEYR 521
DB 498 HCPD--AASTCSTLWCTGSGGLVLCQTKH--FPMADGTPSGEGKMCINGKCVKTRKH 553
QY 522 KRPVVDGGMWPMGSGCSRTGCGVOPSHRECDPPEONGGRVCLGRAYVSGCHTEEC 581
DB 554 PTPPHSGWGMWPMGSGCSRTGCGVOPSHRECDPPEONGGRVCLGRAYVSGCHTEEC 613
QY 582 PP-DGKSPREOQCEKXNAVNTDM-DGNLLQWPKYAGVSPDRCKLFCRARGSEPKYF 639
DB 614 PNNNGTFRREGCEAHNEFSKASFGSGPAVEMIPKAGVSPDRCKLCOAKGIQYFVL 673
QY 640 EAKVIDGTLCGPETALICRGCCVYAGCDHVDVSRKLDKGVCGVCGKNSCRKYSGLTP 699
DB 674 QPKVVDGTPCSTDSVCGQCVKACDRILDSKKFDKCGVCGKNSCTKKSISGVTIS 733
QY 700 TNYGVNDIYTIAGATNIDVKSHPGVNDGNLYALKTADGQVLYLNGNLISAIEDIIL 759
DB 734 ARPGHNDITITETGATNIEVKORNGRSGNNSFLAIKADGTIILNDYTLSTLEQIM 793
QY 760 VEGTILKYSGLATIERLOSFRPLPEPLTVQLLVPGEVFPKVKYTEFVENDVFSNQS 819
DB 794 YKGVAVLYRSGSSAALERIRSFPLKEPLTIQVLTQV-GNALRPKIKYTYFV----- 842
QY 820 SKEPATNTIIOPLHAWTLADMSGCSSTCGAGMORRTVEGCDPSGQASATCNKALXEDP 879
DB 843 KKKKSFNAI-PTPSA-WIIEWGSCSKSCELGMGRVLVECRDINGQASBCEAEVYKAS 900
QY 880 AKPCSQICP 889
DB 901 TRPCADHPCP 910

```

RESULT 2

QSHZMS PRELIMINARY; PRT; 759 AA.

```

AC QSHZMS;
DT 01-MAR-2003 (Trembl) 23, Created
DT 01-MAR-2003 (Trembl) 23, Last sequence update
DT 01-OCT-2003 (Trembl) 25, Last annotation update
DE Metalloprotease (Fragment).
GN ADAMS-1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN 1
RP SEQUENCE FROM N.A.
RA Beorhume D., Russell D.L., Richards J.S., Stirois J.,
RT "Coordinated Regulation of Transcripts Encoding ADAMS-1 and
RT Progesterone Receptor in Equine Preovulatory Follicles";
RL Submitted (AUG-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF541975; AAI1731.1;
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 3.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS50215; ADAM_MPRO; 1.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON TER
SQ SEQUENCE 759 AA; 83410 MW; C1BE4C048918CB9 CRC64;

```

Query Match 43.5%; Score 2112.5; DB 6; Length 759;  
Best Local Similarity 52.6%; Pred. No. 1.2e-167;  
Matches 380; Conservative 131; Mismatches 174; Indels 37; Gaps 13;

```

QY 185 EDESEESOEERAEAGAS-----BPP-----PLGATS-RTKRFVSEARFVETLLVADA 230
DB 1 DDETPGSRGAEPRPGQAGTOMAPRDRAFORAGRPCTGSIKRRKFFSSRYVETMLVADO 60
QY 231 SMAAFYQADIQNHILITMSVAARIYKSPISKINIMVYKVLIVDEBKGPVSDNGILT 290
DB 61 SMAEFHSGKHYLLTLFVAARLYKHPISIRNSVSLVYKVLIVYEBQGPVTSNAULT 120
QY 291 LRNFCNRRFPNCPDRHPHYDTAILLTRNFCQGEGLCDTLGVADIGTICDPNSKCSV 350
DB 121 LRNFCNRRFPNCPDRHPHYDTAILLTRNFCQGEGLCDTLGVADIGTICDPNSKCSV 179
QY 351 IEDEGLQAAHTLAHELGHVLSMPPHDSKPTCTRLFGPMGKHVYAPLFHLNQTLPMSPGS 410
DB 180 IEDEGLQAAHTLAHELGHVLSMPPHDSKPTCTRLFGPMGKHVYAPLFHLNQTLPMSPGS 470
QY 411 AMTLTLLDGHDCCLLDAPGALPLPTGLPGMYALYQLDQCCROQFGPDRPHCEPTSAQ 470
DB 240 AYMITSLFDNGHSECLMDKQNPQLQDGLDPG--TSDANRQOQFTFGESKCHPD--AA 295
QY 471 DVCAQLWC-HTGAEPICHTKNGSLPMADGTPCGPGLCSGSGCLDEEVEYRPPVVDG 529
DB 296 STCTTLMCTGSGGLVLCQTKH--FPMADGTPSGEGKMCINGKCVKTRKHEDTPVHGS 353
QY 530 WAPWGMGSGCSRTGCGVOPSHRECDPPEONGGRVCLGRAYVSGCHTEECPP-DGKSP 588
DB 354 WAPWGMGSGCSRTGCGVOPSHRECDPPEONGGRVCLGRAYVSGCHTEECPP-DGKSP 647
QY 589 REOQCEKXNAVNTDM-DGNLLQWPKYAGVSPDRCKLFCRARGSEPKYFVETLLVADT 647
DB 414 REOQCEAHNEFSKASFGSGPAVEMIPKAGVSPDRCKLCOAKGIQYFVLQPKVVDGT 473

```

```

QY 648 LCGEETLACVRGQCVKAGCDHVVDSPKLDKCGVCGKGNKSGKSLTPNYGNDI 707
D 474 PCSDSDSVSVCGQCVKAGCDRIIDSKKPKKPCICGNGSTCKKIGSTSAKPGVHDI 533
QY 708 VTIPAGATNIDVKORSHPGVQNDGNVYALKTADQOYLINGNLASIAIEDILVKGITLKY 767
D 534 VTIPGATNINEVKORNGRSHNGSFLAKADGTIILNGDFTLSTLEQDITVKGSLRY 593
QY 768 SGSTATLERLQSPFLPEPLTVOLLTVGGEVFPFKVYTFPVNDVDFSMQSSKERATN 827
D 594 SGSSAALERIRSFSPLEPLTVOLLTVGNAARPKIYTYFV-----KKKESFN 642
QY 828 IIPQLHAQWYLGDMSECSSTCGAQWQRRVYECRDPESQASATCNKALKEEDAKFCESOL 887
D 643 AIPF--SEWVIEEWGECSSKCGQWQRRVYECRDINGQPASECAPKASTPCADLP 700
QY 888 CP 889
D 701 CP 702

RESULT 3
Q8K384 PRELIMINARY; PRT; 833 AA.
ID Q8K384
AC Q8K384
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Similar to a disintegrin-like and metalloprotease (repolysin type)
DE with thrombospondin type 1 motif, 4.
GN ADAMTS4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Straube R.;
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027773; AAH27773.1; -.
DR MGD; MGI:133949; Adamts4.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_Zn_Bs.
DR InterPro; IPR008884; TSP1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; tep_1; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1.
DR PROSITE; PS00215; ADAM_MERO; 1.
DR PROSITE; PS00092; TSP1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR InterPro; Protease; Metalloprotease.
KM Integrin; Protease; Metalloprotease.
SQ SEQUENCE 833 AA; 90097 MW; 4CEC83DFEC3AA619 CRC64;

Query Match 37.4%; Score 1813; DB 11; Length 833;
Best Local Similarity 45.0%; Pred. No. 1.66-142;
Matches 370; Conservative 129; Mismatches 247; Indels 76; Gaps 23;

```

```

QY 126 SFLDDEEFTIOP-----QGAGSLAOPHRLOKRWGPARGARLPFGPEWEVETGEGQROE 179
D 148 VQYGAELHLLOPLLGALNSAGPGA--HILRRSPAS---OGPMCTVAPSGS--- 198
QY 180 RGDHODESEESQEEBAGASEPPLIGATSTRTKEFVSEARFETLLVADASMAAFYGAD 239
D 199 -----PSPI--SRTRKFPASIREVELTVVADDDGAAPHGTG 233
QY 240 LQNHILLMSVARIYKHPISINSLMWVYKLYIEDSKWGEVENDGSLTLRNQNMOR 299
D 234 LKRYLTVMAAARAKFKHPSIRNPVNLVYTRLVILGSGEGEQVPSAAQTLRSCTWOK 293
QY 300 RENOPSDRPEHYDTAIIITRONFCQGBGLCDTLGVADIGTICDPNKSQSVIEDGELOAA 359
D 294 GLNTPNDSDPHFTAILFTTRQDLG-VSTCDTLGMADVGTCDPARSCAIVEDGLQSA 352
QY 360 HTLAHELGHVLSMPDDSKPCTRLFGPMG-KHHNMAPLVHNLQTLPMSPGAMLTLL 418
D 353 FTAHHLGHVFMMLHDNSKPCNTNLNGQGSRRHVA PVAAHVDPPEWSPCSARFTTFL 412
QY 419 DGAHGDCLDAPGAALPLFTGLPGRMALYOLDQCCROIFGDPFRHCPTNSAQDVCAOLMC 478
D 413 DMGYGHCLLDKPEALHLPATFPGKD--YDADROQLTFGPDSSHPQLPP--CALMC 468
QY 479 --HTDGAEPYCHTKNGSLFMADGTCGPRHLCSBSCLPBEVEVERPKVYDGGALPMGFW 536
D 469 SCHLNG-HAMCQTKIS--FMADGTPCGSSQA CMGRCCLHVDLKD FNPVQAQGMGPWGW 525
QY 537 GECSTRCGGVQFHSRECKDPEPONGRCYCLGARAYKOSCHTEPCP-PDGKFFRQOCEK 595
D 526 GDCSRICGGGVQFSSRDCRFPVRNGKCYCEBRKRRFRSCTNENPHOSALTFRREQCA 585
QY 596 YNAVNTYMDGNL--LQWPKYAGVSPDRCKLFCRARGSEFVFAKYVDITGLCGPE 652
D 586 YN--HRTDLEFKSPPEPMVMVPRYTGVAAPRDOCKLTCQARALGYVYVLEPRVADGTPCCPD 643
QY 653 TLACVARGQCVKAGCDHVVDSPKLDKCGVCGKGNKSGKSLTPNYGNDIVTIPA 712
D 644 TSSVCVQGRCHAGCDRIIGSKKPKKPCICGNGSTCKKIGSTSAKPGVHDI 703
QY 713 GATNIDVKORSHPGVQNDGNVYALKTADQOYLINGNLASIAIEDILVKGITLKYSGSI 771
D 704 GATHILVRQGGSGGLKS--IYALATLSDGSVALNGEYILMSPSTVLPVGAVALRSYGAT 761
QY 772 ATLRLQSPFLPEPLTVOLLTVGGEVFPFKVYTFPVNDV 813
D 762 AASETLGSHGLAQLTVOLLTVAGNPONARLRSFFVPRPV 802

RESULT 4
Q8BNJ2 PRELIMINARY; PRT; 845 AA.
ID Q8BNJ2
AC Q8BNJ2
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE A disintegrin-like and metalloprotease.
GN ADAMTS4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Body;
RC MEDLINE=22354683; PubMed=12466851;
RA The RANKM Consortium;
RA The RANKM Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).
EMBL; AK083534; BAC38944.1; -.

```

DR MGD; MG1:1339949; Adamt4.  
 DR GO; GO:0004222; F:metallopeptidase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR006586; ADAM cysteine.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01421; Repolysin; 1.  
 DR Pfam; PF00090; tsp 1; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00608; ACR; 1.  
 DR SMART; SM00209; TSP1; 1.  
 DR PROSITE; PS50215; ADAM\_PROPRO; 1.  
 DR PROSITE; PS50092; TSP1; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR SEQUENCE 845 AA; 91239 MW; B496C3190D1A9225 CRC64;

Query Match 37.4%; Score 1813; DB 11; Length 845;  
 Best Local Similarity 45.0%; Pred. No. 1.7e-142;  
 Matches 370; Conservative 129; Mismatches 247; Indels 76; Gaps 23;

15 LLLLLLLPLARGARPARPAAGQASELVVPTRLPGSA-----GELALHLSAFGRGV 66  
 |||||  
 46 LLLLLLFLSLA--WPASFLP--REBEIVFPEKLNSSILPGSGVPAARLLYRLPAFGEMTL 101  
 |||||  
 67 LRLADDSFLAPERFIERLGGSGRATGGERGLRGCFBGTGYNGBESLAVALCRG--LSG 125  
 |||||  
 102 LELEDPGVQVEGLTVQYLQAPENLQGA--PGYILGTGNGPESVSLHMDGALLG 159  
 |||||  
 126 SFLLDGESEFTIOP-----QGAGSLAQPHLQRMGPAPARLPDRCGEWEYETGEQROE 179  
 |||||  
 160 VLYQRGALHLQPLEGALNSAGPGA--HILRRKSPAS-----QGMCTYKASGS--- 210  
 |||||  
 180 RGDHODEBESQEBEAGASPPPLGATSRTRKFSVEARFVETLLVADASMAAFGAD 239  
 |||||  
 211 -----PSPF--SRTRKFSLSRFVETLLVADAKMAAFRG 245  
 |||||  
 240 LQNHILTMSVARIYKPSIKNSINLMVVKYLVEDEKMGPEVSDNGSLTRNFCMOR 299  
 |||||  
 246 LKRYLLTMAAAAKAFKPSIRNPVNLVTRVLVILGSGBGPQYGPBAQTLRSFCWQR 305  
 |||||  
 300 RPNQSDHPEHYTALILTRQNFQSGEGCLDTIGVADIGTICDPNKSQVIEDEGLQAA 359  
 |||||  
 306 GLNTPNDSDPHFDAILFTRODLG--VSTCDTLGMADVGVCDPASCAIVEDDGLQSA 364  
 |||||  
 360 HTLAHELGHVLSMPHDSKPCRLFGPMWG--KHNYMAPLFVHNLQTLFMSPCSAMYLTEL 418  
 |||||  
 365 FTAAHELGHVFNMLHDKPCTNLNGQSSRAHWAPMAHVDPEEPMSCSAFTIDFL 424  
 |||||  
 419 DGHGDCLLDAPGAALPLPTGLPGEMALYQLDQCRQIFGDFRHCNTSAQDVCAQWMC 478  
 |||||  
 425 DNGYHCHCLLDPEAPLHLPTATPGKD--YDADROCOLTFGSDSHCPQLPPP--CALIWC 480  
 |||||  
 479 --HTDGAEPRLCHTKNGSLPMAADGTPCGPGLHCSGSCLPBEEVERPKPVVDGGAAPMGW 536  
 |||||  
 481 SGHLWG--HAMCQTKS--PMADGTPCGSSQACMGARCLHVDQLQDFNVPAQAGGPMGPW 537  
 |||||  
 537 GECRTCCGGVQFHSRECKDEPONGRGYCLGRBAKYQSCHEBEP--PDGKSFREQCCEK 595  
 |||||  
 538 GDCSRTCGGVQFSSRDCTRPVPRNGKCYCEGRTRFRSCNTENCPSGSLATFEEGCA 597  
 |||||  
 536 YNANYTMDGNL--LOWPKYAGVSRDRCKPCRAAGSRKRYEAKYIDGTLGPE 652  
 |||||  
 538 YN--HRTLFKSPGPMWVPRYTGVARDDCKLTCQARALGYVYVLEPRVADGTPGSPD 655  
 |||||  
 653 TLATCVRGQCVYAGCDHVVDSPRLDKGCVGCGKSGKSKYSGSLTPTNYGYNDIVIPA 712  
 |||||  
 656 TSSVVGQRCHTAGGDRITGSKKFKDKCMVCGGSGKSGSKSGSPFKFRGYSVYVIPA 715  
 |||||  
 713 GATNIDVKQRSHPGVQNDGNTALKTADGQYLLNGNLAIISAIEDDILVKGTT--LKSGS 771  
 |||||

Db 716 GATHILVROQGGSLKS--ITYALKLSDSVALNGEYTLMPSPDVLPGAVSLRYSGAT 773  
 |||||  
 Qy 772 ATLERLOSFPPLPEPLTVQLTVPGEVFPFKYTFEYNDV 813  
 |||||  
 Db 774 ASETLSGHPPLAQPLITLQVL--VAGNPQARLRYSPFVRPV 814  
 |||||

## RESULT 5

Q7YS95 PRELIMINARY; PRT; 839 AA.  
 AC Q7YS95;  
 DT 01-OCT-2003 (T-EMBLrel. 25, Created)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)  
 DE Aggrecaenae-1.  
 GN ADAMTS-4.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 OC Bovidae; Bovinae; Bos.  
 CX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Arai M., Anderson D., Annis B., Collins-Racie L., Corcoran C.,  
 RA Di Blasio-Smith E., Morris E., Dörner A., Lavallie E.,  
 RT "Cloning and characterization of bovine aggrecaenae-1."  
 DR Submitted (May-2002) to the EMBL/Genbank/DBJ databases.  
 DR EMBL:AB516915; AAF47196.1; --  
 SQ SEQUENCE 839 AA; 90279 MW; 1B488A27DF5B96B1 CRC64;

Query Match 37.0%; Score 1794.5; DB 6; Length 839;  
 Best Local Similarity 45.0%; Pred. No. 5.8e-141;  
 Matches 367; Conservative 124; Mismatches 254; Indels 71; Gaps 22;

15 LLLLLLLPLARGARPARPAAGQASELVVPTRLPGSA-----GELALHLSAFGRGV 67  
 |||||  
 39 LLLLLLFLPSA--WPASFLP--REBEIVFPEKLNSSILPGGAPARLLYRLPAFGEMTL 94  
 |||||  
 68 RLAPDPSFLAPERFIERLGGSGRATGGERGLRGCFBGTGYNGBESLAVALCRG--LSGS 126  
 |||||  
 95 ELEKDPGVQVEGLTVQYLGRAPELLQGA--PGYILGTGINDPESVSLHMDGALLGV 152  
 |||||  
 127 FLIDGSEFTIOP--QGAGSLAQ--HILQRMGPAPARLPDRCGEWEYETGEQROE 182  
 |||||  
 153 LQYRGTELHIQPLEGAPNSAGPANHILRRKSPSG-----QGMCTYKASGS--- 197  
 |||||  
 183 HQDESEBESQEBEAGASPPPLGATSRTRKFSVEARFVETLLVADASMAAFYADIQ 242  
 |||||  
 198 -----APRGRSPSPRAKRFASLSRFVETLLVADAKMAAFHAGLKR 240  
 |||||  
 243 HILTMSVARIYKPSIKNSINLMVVKYLVEDEKMGPEVSDNGSLTRNFCMORFN 302  
 |||||  
 241 YLLTMAAAAKAFKPSIRNPVNLVTRVLVILGSGBGPQYGPBAQTLRSFCWQR 300  
 |||||  
 303 QPSDHPHYTALILTRQNFQSGEGCLDTIGVADIGTICDPNKSQVIEDEGLQAAATL 362  
 |||||  
 301 TPDDADPGHFDAILFTRODLG--VSTCDTLGMADVGVCDPASCAIVEDDGLQSAFTA 359  
 |||||  
 363 AHELGHVLSMPHDSKPCRLFGPMWG--KHNYMAPLFVHNLQTLFMSPCSAMYLTELDDG 421  
 |||||  
 360 AHELGHVFSMLHDKPCTNLNGQSSRAHWAPMAHVDPEEPMSCSAFTIDFLNG 419  
 |||||  
 422 HGDCLLDAPGAALPLPTGLPGEMALYQLDQCRQIFGDFRHCNTSAQDVCAQWMC--H 479  
 |||||  
 420 FGHCLLDPEAPLHLPTATPGKD--YDADROCOLTFGSDSHCPQLPPP--CALIWC 475  
 |||||  
 480 TDGAEPRLCHTKNGSLPMAADGTPCGPGLHCSGSCLPBEEVERPKPVVDGGAAPMGWEC 539  
 |||||  
 476 LMG--HAMCQTKS--PMADGTPCGPAQACMGARCLHVDQLQAFNVPAQAGGPMGWSGDC 532  
 |||||  
 540 SRTCCGGVQFHSRECKDEPONGRGYCLGRBAKYQSCHEBEP--DGKSFREQCCEKRNA 598  
 |||||  
 533 SRSCTGGVQFSSRDCTRPVPRNGKCYCEGRTRFRSCNTDCTPSALTFRBBQCAAYN-- 591  
 |||||

```

QY 599 XNYTMDGNL---LOWPKYAGVSPRDRCKLFCARGRSEKFEAKYIDGTLGPETLA 655
DB 592 -HRDLEFNFGPMWMDVRYGVAPRDCKLTCQTRALGYGVLDPRVADGTPCSPDSS 650
QY 656 ICVRQACVKGAGCDHVVDPRKLDKGVCGGKSGKSGKRVSGSLPTNYGYNDIVTIPAGAT 715
DB 651 VCVRGRCCHAGCDRTIGSKKFKDKCMVCGGSSCSKSGSFKFRGYNVVTIPAGAT 710
QY 716 NIDVKQSHPGVQNDGNLALKTADGQYLNGNLAIISAIEDIIVKGTI-LKYSGIATL 774
DB 711 HILVACQSPSVRS--LYIALKLPGSYALNGEYTLIPSPDVLPAGAVSLRYSATTAAS 768
QY 775 ERLQSFRLPERLTIVQLTVPGEVPPKVKYTFVP 810
DB 769 ETLGSGHPLAEPLTQVL-VAGNPNONARLRYSPFP 803

RESULT 6
Q8BGP4 PRELIMINARY: PRT; 623 AA.
AC Q8BGP4;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE A disintegrin-like and metalloprotease (Fragment).
OS ADAMTS5.
OC Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; Tissue=Cerebellum, and Retina;
RA MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RT Nature 420:563-573 (2002).
RL EMBL; AK044746; BAC38062.1;
DR EMBL; AK044746; BAC38062.1;
DR MGD; MGI:1346321; Adamts5.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008318; F:protein prenyltransferase activity; IEA.
DR GO; GO:0018270; F:zinc ion binding; IEA.
DR GO; GO:0018346; P:protein amino acid prenylation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_N_Zn_BS.
DR InterPro; IPR002088; PPA.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR Pfam; PF01421; Reprolysin; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; ACR; 1.
DR SMART; SM00209; TSP1; 2.
DR PROSITE; PS00215; ADAM_MPRO; 1.
DR PROSITE; PS00904; PPA; 1.
DR PROSITE; PS0092; TSP1; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON_TER
SQ SEQUENCE 623 AA; 68681 MW; 80CE9BCDACA37502 CRC64;
Query Match 30.6%; Score 1483.5; DB 11; Length 623;
Best Local Similarity 44.7%; Pred. No. 4e-115;
Matches 284; Conservative 93; Mismatches 238; Indels 21; Gaps 10;
QY 261 KNSINLMVYKVLIVDEKMGPEVSDNGSLTIRNFCMQRFRNPQSDRHPHYDTAILLTR 320
DB 1 ENHRLAVVYKVVLTDRDTSLEVSNNATTLTKNCKQOHQDGHDDHEHYDAALIFTR 60

```

```

QY 321 QNFCGEGLCDTLGVADIGTICDPNKSQSVIEDSGIQAAHTLAHELGHVLSMHPHDSKFC 380
DB 61 EDLCGHHS-CDTLGMADVGTICSPERSCAVIDDDGHAFAVHAHEIGHLIGSHDSKFC 119
QY 381 TILFGMGHGMARPLFVHLNQTLPMSPCSAHYLTLELDGGGDDCLDAPGAALPLPTEL 440
DB 120 EENFGTTEKRLMSLTLSIDASKWSKTSATITFEFLDDGKGLDLPRQIILGPEEL 179
QY 441 PGRMALYOLDQOCROQFGPDRHCPNTSAQDYCAQJWCH-TDGAELCHTKXGSLPMADG 499
DB 180 PGQ--TYDATQCCNLTFGEYSVCP---GMDYCARLCAVVRQGVNCLTK--KLPAVYG 232
QY 500 TPQGPCHLSSESCLPEEVERPKPVVGGMALPMKPMGECSTTCGGVGFHSRECKDPER 559
DB 233 TPQGRVCLQGGCDVTKKXKTYSTSHGNMWSMGQCSRSQGGVGFAYRHCNNPAP 292
QY 560 QNGRYCLGRAPKAYQSCHTEECPDQKSPREOCCEKYNAYNTMDG--NLQWVPKYAG 617
DB 293 RNSGRCTGRALVYNSCSVTPCPBNKSPRHQCAKNGYQ--SDAKGYKTFVEMVPKYAG 351
QY 618 VSPDRCKLFCARGRSEKFEAKYIDGTLGPETLAICVGGCYKACDHYVDSPRKL 677
DB 352 VLPADCKLTCRAKGYGVVSPKVTDTGCRPYNSNCVARGCRVTCGDGIIGSKLYQ 411
QY 678 DKGVCGGKSGKSKRVSGSLPTNYGYNDIVTIPAGATNIDVKQSHPGVQNDGNVLA 737
DB 412 DKGVCGGKSGKSKTKIIGTFNKSQGYDVAVIPGATHIKVRQFRAKQGTFTAYLALK 471
QY 738 TADGQYLNGNLAIISAIEDIIVKGTILKYSIATLEPL--QSPRLPEPLTIVQLTVP 795
DB 472 KKTGEYLINGKYMISFETIIDINGVYVNSGMSHRDPLHMGYSATREIIVQLTAD 531
QY 796 GEVFPK--VYTFPVENDVDVFSMQSKERATVNIQILHAQVLTGPMSCSSSTCGAG 852
DB 532 ---PTKALDVRYSPFPVKTKTKQKNSVSHSNKVGPRHSTOLQWTVGMLACSRICDTG 587
QY 853 MQRRTVECDPSGQASATCNKALKEADAKPCESQJC 888
DB 588 MHTRTVQCQDGNRKLAKGLLSQPSAFKQCLIKKC 623

RESULT 7
Q19791 PRELIMINARY: PRT; 2165 AA.
AC Q19791; Q27524;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE F25H8.3 Protein.
GN F25H8.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gajadasy S.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Almscough R., Anderson K., Baynes C., Berke M., Boulfield J., Burton J., Connell M., Copesey T., Cooper J., Coulson A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd J., Mcmurtry A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R., Shalton N., Smith A., Sonhammer E., Staden R., Sulston J., Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sprat J., Wohlman P.;
RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

```

RT elegans. ;  
 RL Nature 368:32-38 (1994).  
 DR EMBL; 269361; CAA93288.1; .  
 DR EMBL; 269360; CAA93288.1; JOINED.  
 DR EMBL; 269360; CAA93287.1; .  
 DR EMBL; 269361; CAA93287.1; JOINED.  
 DR F1R; 121371; 121371.  
 DR HSP; P15167; 1DTH.  
 DR MEROPS; M12.135; .  
 DR WormPep; F25H8.3; CE05729.  
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:000508; F:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001590; Peptidase\_M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR000884; TSP1.  
 DR Pfam; PF01421; Reprolysin; 1.  
 DR Pfam; PF00090; tsp 1; 14.  
 DR SMART; SM00209; TSP1; 18.  
 DR PROSITE; PS00215; ADAM\_MEROPS; 1.  
 DR PROSITE; PS00092; TSP1; 15.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 DR SEQUENCE 2165 AA; 244397 MW; FCC3DABA9CA888 CRC64;

Query Match 30.4%; Score 1476; DB 5; Length 2165;  
 Best Local Similarity 34.1%; Pred No. 1,1e-113;  
 Matches 327; Conservative 150; Mismatches 319; Indels 164; Gaps 29;

46 RLPGSAGELALHLSAFAKGVLRRLAPDDSL-----APEFKIERLQGS 88  
 DB RLQGVARDG-----GHACHRLRLSDDAVYVHLRWNNQIPDSHNKSVPHFSNFPAM 148  
 QY 96 GRATGGERGLRG-----CFPSGTNAGEPSLAVALC---KGLSFLDDEHPTIQ 137  
 DB 149 VLYLDESEERVGMSRTDPDCTYRAHVGVQH-SIVNLCSDEGLYMLPFGIHIVE 207  
 QY 138 P-----QAGAGSLAOPHRLRQWGPAGARPLFRGPEVEVEGEGQERGDH-CEQSEE 189  
 DB 208 PLSNGTNEHDGASHRQHLVAKKPDPMFKSF-----DHLSNYSVN 248  
 QY 190 ESQSEBAGASPPPLGATSTRTKPV-SEARFVETLLVAASAAPGADLQNHILTM 248  
 DB 249 ETEETVAATWQDQMEDVIERKASRRANSMWHYEVLVADTKYETVGRSLERYVTLF 308  
 QY 249 STAAATYKHPISIKNSINLMVYVLLVEDEKWPGEVDNAGLTLNFCWQRRFPQSPDRH 308  
 DB 309 STVASTYKHSIRASINNVVVLIVLKTENAGFRITQNAOQLQDFCMQGYNDPDDSS 368  
 QY 309 PEHYDTAILTRQNTQGOEGELCDLVGADIGTICDPNKSQSVIEDEGLQAHTLAHELGH 368  
 DB 369 VOHHVVALITRKDCRSQKCDTGLAELGTMCQMSKCAIIEINGLSAFTTAHEIGH 428  
 QY 369 VLSMHDSPKCTRLRFGPMGK-----HYMAPLFYVLTNTLWSPSCSA 411  
 DB 429 VFSIHDERKKS-TYMPVNNKCKFQSTKFDKQNFHMAPLFLENTNPMNSPSCSA 487  
 QY 412 MYLTELDDGNG--DCLLDAPGAL---PLPTGPMALYQLDQCRQIFGDFRHP 465  
 DB 488 GWLEFFLENNRGQTCLFQDQVERRYEDVFRDEPGKK-YDAHQCKFVGASELCP 545  
 QY 466 NTSADVCAQALWCHN-DGAEPLCHRNKNSLPMADSTPGPGH--LCSGSL---PEEV 519  
 DB 546 ---YMPCTRLMCAITFYSGNGCRITQH--MPADSTPCDEBSRSMFCHAGACVRLAPESLT 600  
 QY 520 EEPKPVVDGVAFWPWPWECSTFGCGVQVFSHRECKDEPONGRGYCLGRBAKYQSCHTE 579  
 DB 601 K-----IDQWGDMSWEGSCSTFGCGVQVKGRLDSCSPKPNRGKXCVGQERYRSNTQ 655  
 QY 580 EEPPOGKSPREOGCKEKN--ANNYDMNGNLQWPKYAGVSPRRCLCFRAGRBSFK 637  
 DB 656 EEPWQTPQREVQCEPNKNDIGIGIVASTNTHWPKYANVAPEKRLCYRLSGSAFY 715  
 QY 638 VFEAKVIGTLGPELTALCVAGQCVKAGCDHVVDSPKLDKCGVCGKSGNSCKRVGSL 697

DB 716 LMDKTVDDGTPCORNDDICVAGACWPGACDHLHSTRDKCGVCGGDDSCCKVKGTF 775  
 QY 698 TPT-NGYNDIVITIPAGANINIDYKQSHPGVQNDGYTLAKPDAGQYTLNGNATIAIRQ 756  
 DB 776 NEQGTGYEWMKIPGASNIDIRQGYNNKEDDYLSTRANGEFLLNGHFQVSLARQ 835  
 QY 757 DILVKGITLKYSGSIATLRLQSFRLPEPLTVQLTVGEVFPKRYTFF--VPNDV 813  
 DB 836 QIAFQDPTVLEYSSDILIRINGTPIRSDIVYHLVSQSH--PPIISYEVYTAAPNAV 893  
 QY 814 -----DFSN-QSKERATNTIQL----- 832  
 DB 894 IRPISALYLMRYDTWTEDCRACRQOQSKLMCLDMSHRSQSHRNCQNVKPKQATRM 953  
 QY 833 ---LHAQWVLGDMSECSSTCGAGMQRFTVEGRDPSCG--QASAT---CNKALKPEDAKPC 883  
 DB 954 CNIDCSTRTWTEDEVSSCAKCGSGQKRGKRVSCVTKMGDRQTAASEHLCDRNSKPSDIASC 1013

## RESULT 8

Q8SX80 PRELIMINARY; PRT; 1688 AA.

AC Q8SX80; Q9VP61;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE GH16393P (CG14869 protein).  
 GN CG6107.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;  
 OC Tephritidae; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=Berkely;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Chame M., Chavez C., Dorsett V., Dresner D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guertel H., Kronmiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nunco J., Pacled J., Pargass V., Park S.,  
 RA Patel S., Phouenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Ceiniker S.,  
 RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.

RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=Berkely;  
 RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amandides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Chame M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt R.G., Nelson C.R., McKlos G.L.G.,  
 RA Arril J.F., Agbayani A., An H.-J., Andrews-Piankocch C., Baldwin D.,  
 RA Bailey R.M., Basu A., Baxendale J., Beyraktaroglu B., Beasley E.M.,  
 RA Beeson R.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,  
 RA Bortolova D., Botchan M.R., Bouck J., Brokstein P., Bottler P.,  
 RA Butts K.C., Busan M.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Domes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glaeser K.,  
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalili M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kratz C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li X., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Moberly C., Morris G., Mostrel A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclel J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Stien-Klamas I., Stapleton M., Skupski M.P., Smith T.,  
RA Spier E., Spirdling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu Q., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Celniker S.E., Adams M.D., Krommler B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Barzon J., An H., Baldwin D., Banston J., Beeson K.Y., Bissam D.A.,  
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Fartan D.,  
RA Ferreira S., Frisbe E., Galle R.F., Gang N.S., George R.A.,  
RA Gonzalez M., Houch J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ingwen C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
RA Paclel J., Parages V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puti V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,  
RT "Sequencing of *Drosophila melanogaster* genome.";  
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Maira S., Crosby M.A., Matthews B.B., Bayraktaroglu I., Campbell K.,  
RA Hixey P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Bernan B., Carlson J.W., Celniker S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frisbe E., de Grey A., Harris N.,  
RA Krommler B., Marshall B., Milburn G., Richter J., Russo S.,  
RA Searle S.M.J., Smith E., Shu S., Smutnick F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,  
RT "Annotation of *Drosophila melanogaster* genome.";  
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.,  
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RL Submitted (SEP-2002) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AY094716; AAM11069.1; -;  
DR EMBL; AE003709; AAF5199.2; -;  
DR FlyBase; FBgn0038340; CG6107.  
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
DR GO; GO:0008270; F:zinc ion binding; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001590; peptidase\_M12B.  
DR InterPro; IPR002870; peptidase\_M12B.N.  
DR InterPro; IPR006025; pept\_M12B.  
DR InterPro; IPR000884; TSP1.  
DR Pfam; PF01562; Peg\_M12B\_piroged; 1.  
DR Pfam; PF01421; Reprolysin; 1.  
DR SMART; SMO0209; TSP1; 5.  
DR PROSITE; PSS0215; ADAM\_MEPRO; 1.  
DR PROSITE; PSS0092; TSP1; 3.  
DR PROSITE; PSS0142; ZINC\_PROTEASE; 1.  
SQ SEQUENCE 1688 AA; 189867 MW; 48FB8DD4DE0CA4D2 CRC64;  
Query Match 24.7%; Score 1198; DB 5; Length 1688;  
Best Local Similarity 30.0%; Pred. No. 1.4e-90;  
Matches 326; Conservative 144; Mismatches 343; Indels 274; Gaps 40;  
23 PLARGAPRPA-----GGQASLVVTR----- 46

Dd	197	PLANSPPATPRTLEPRRLDNDTSTDHPPDDGLDDEEHSFVNFPTKYVNSISEADI	256
Qy	47	-----LPQSAGEL-----	76
Dd	257	YESKRNSDINSFLKESASAFMTGTYRMSNEIMDPHQYMLNVFGRLHVLRODASFV	316
Qy	77	----APEKI-----ERLGGSGRAIGGERGLRGCFSGTYNBPESLAAVLCRGLSGSF	127
Dd	317	HNSWTHIRILKEGEHPPTTEABEABRHL-GCYSGVEDDEPMSVSLCGGTGTI	375
Qy	128	LJLDEEFTIQOGAGSGLAOPHRLOPWPAGARPLPRPEWEVEVEGEGORQERG-DHOD	186
Dd	376	KTSFGALLIQVNNTSDEVLRHVPRKSQRABHNV-----SKPELGDDEMS	423
Qy	187	SEESQEEAAGASEPPPLGATSITTRFV-----SEAFVETLLVADMSMAFYADLON	242
Dd	424	KLEQVQEEQSKSR-----KLNRKKKHVADVNDQVYTLLEVLAVDMSKQFHEDDLP	477
Qy	243	HLITMSVAAIRYKHSPIKSNINMLVAYEDEKMWPEYD--NGGLTRNFENMQR	299
Dd	478	YTLIMSVSSIFPDASIGNSIRLLVRLISL-----PNNDQHSNEMLKHQCOF--	529
Qy	300	RFNQPSDBRPHDYTALILTRQNTFCQ--EGJCDTLGVADIGTICDPNKSQSVIEDBGLQ	357
Dd	530	-INOSGYER---DTAMLITREPICGSVPRKICMHLGLAEIGTVCS-SSSCSIVDDTGLP	583
Qy	358	AAHLTAHLGHVLSMPHDSKPC---"TLRFGPMKHVMAPLFHLNQTLEPMSPCSAHY	413
Dd	584	TALTAHLGHITLNMNHDDDKCPYTRQNNKXVLAHMSWGIMH-PMWSKCSRF	642
Qy	414	LTELJLDDGHPDCLLDARPALPLPT-GLPRMALYUOLDQOCROIIPGDFRHCNPNTSAOV	472
Dd	643	VSEFLKTKDSCLEFSGVAHIFYGERLPGE--IYSIDAQQLSFGNDFGFCP--TDEE	697
Qy	473	CAQIMCH-TDG-APRLCHTKXGSLPMADGTFCG--FGHCSGSGCLPEEVEVERPKVYVNG	529
Dd	698	CKRLMCRKTSNSNEQCASN--LPMADGTGCGSGHMCQKVCNSHGVRQ--VWGG	753
Qy	530	WAPGPMQECRGTGGGVOFGSHRECKPEFONGRVCILGRBAKYCSCTHECPRPDGSFR	589
Dd	754	WGPTPLFTGSLTGCGVQESRRCONQPVENGDKYCTGSKXKYSRCSNTHGCPSPMPR	813
Qy	590	EQCEKTAANYTMDGNL-LQNVFKTAGVSPRDCXLFCSARGRSEKVFBEAKYIDGTL	648
Dd	814	EQCCYAMGRMNIIPGVNPDTKWPKY---EKDACKLFCGMDMKVTFEMLKSMTDGTG	869
Qy	649	CGEFTLAIYVRGOCVVKACDHVDSPRLLDKCGVGGGSGNSCRKYSGLTPTN-YGYND-	706
Dd	870	CAVDSFDKCVNGICRPACDNEHLSIAKLDGCVGBGANDTCHETVGNLVLNSLGLDNG	929
Qy	707	-----IYTIAGATNIDVKORSHPGVNDENTYALAKTADQYLLNGMLAISABOD	757
Dd	930	NEPMKTYUYVRIIPKGSANIITITQRPV---DQNFIVLTDROBELLNGKF-LKTYPLK	984
Qy	758	ILVKGTLIKYSGSIALERLQSF--PLREBLTYOLLVREVEPRK-----YKTEFPV	809
Dd	985	FVAGVTMOTQYSSVVEQVNTTYSMKLSRDLVQIITSL-DVBSKQODTVLISYSTTI	1044
Qy	810	PNDVDF-----	821
Dd	1043	DKPRDYAEVETIYRWMQAPENCDSLCEGRSHRLPACISTTQGVKVARQFCDSAMPKID	1100
Qy	822	ERA-----TNIIOPLHAQ-----	836
Dd	1103	DRACNTDORLNLVYTSISECSACGELGTREKTYACQVTFNMPRNSNIVMSYCKLADV	1167
Qy	837	-----WVLGDWEGCSGTGAGMORATVECRDPSGQAS-ATCNKALPE--DAKP	882
Dd	1163	AYHECRGECGVLSWSTGSKSCGTGQONEAHCYLHNSRVSDDLONPRTPHLNTLLGI	1222
Qy	883	CESQLCP 889	



```

DB      1223 CNTSCOP 1229

RESULT 9
O9M493  PRELIMINARY; PRT: 1059 AA.

ID      O9M493;
AC      O9M493;
DT      01-MAY-2000 (TRENBLREL. 13, Created)
DT      01-OCT-2002 (TRENBLREL. 22, Last sequence update)
DT      01-OCT-2003 (TRENBLREL. 25, Last annotation update)
DE      CG4096 protein.
GN      CG4096.
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Peerygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_TaxID=7227;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Berkeley;
RX      MEDLINE=20196006; PubMed=10731132;
RA      Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA      Sutton R.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA      Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA      Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA      Abiri U.F., Agbayani A., An H.-U., Andrews-Pfannkoch C., Baldwin D.,
RA      Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA      Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolintsov S.,
RA      Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA      Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA      Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA      de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA      Dodson K., Dopp L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA      Durbin K.U., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA      Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA      Gload A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA      Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
RA      Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbegan C.,
RA      Jafari M., Kalush F., Karpen G.H., Ke Z., Kennison U.A., Ketchum K.A.,
RA      Kimmel B.E., Kodali C.D., Kratz C., Kravitz S., Kulp D., Lai X.,
RA      Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA      Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA      Merulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,
RA      Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA      Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA      Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA      Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA      Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA      Spier E., Spirding A.C., Stapleton M., Strong R., Sun E.,
RA      Svitzkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA      Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA      Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA      Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA      Zheng R.A., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA      Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT      "The genome sequence of Drosophila melanogaster.";
RL      Science 287:2185-2195 (2000).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RC      Evans C.A., Gocayne J.D., Amaratilake P.G., Brandon R.C., Rogers Y.,
RC      Barzon J., An H., Baldwin D., Barzon J., Beeson K.Y., Busan D.A.,
RC      Carlson J., Center A., Champe M., Davenport L.B., Dietz S.M.,
RC      Dodson K., Dorsett V., Dopp L.E., Doyle C., Drenek D., Farfan D.,
RC      Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RC      Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RC      Ibbegan C., Jafari M., Kruse D., Li P., Mattei B., Moshrefi A.,
RC      Moshrefi A., Moy M., Murphy B., Murphy L., Nelson C.R., Nuno J.,
RC      Pacle J., Parag V., Park S., Patel S., Pfeiffer B., Scheeler F.,
RC      Shounanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RC      Stapleton M., Strong R., Svitzkas R., Tector C., Tyler D.,

```

```

RA      Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT      "Sequencing of Drosophila melanogaster genome.";
RL      Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RC      Maira S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RC      Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RC      Tupy J.L., Bergman C., Bernan B., Carlson J.W., Celniker S.E.,
RC      Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RC      Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RC      Seale S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RC      Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT      "Annotation of Drosophila melanogaster genome.";
RL      Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE FROM N.A.
RC      Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter J.C.;
RL      Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN      [5]
RP      SEQUENCE FROM N.A.
RC      Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RL      EMBL; AB003435; AAFA6065.2; -
RC      HSSP; P15167; 1ATL.
RC      MEROPS; M12.231; -.
DR      GO; GO:0004222; Fimetalloendopeptidase activity; IEA.
DR      GO; GO:0008270; Zinc ion binding; IEA.
DR      GO; GO:0006058; Proteolysis and peptidolysis; IEA.
DR      InterPro; IPR001590; Peptidase M12B_N.
DR      InterPro; IPR002870; Peptidase M12B_N.
DR      InterPro; IPR000884; TSP1_M_zn_BS.
DR      InterPro; IPR008085; TSP1.
DR      Pfam; PF01562; pep_M12B_propep; 1.
DR      Pfam; PF01421; Reprolysin; 1.
DR      Pfam; PF00090; tsp_1; 2.
DR      PRINTS; PR01705; TSP1REPEAT.
DR      SMART; SM00209; TSP1; 2.
DR      PROSITE; PS50215; ADAM_MEPRO; 1.
DR      PROSITE; PS50092; TSP1; 2.
DR      PROSITE; PS00142; ZINC_PROTEASE; 1.
DR      PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ      SEQUENCE 1059 AA; 119195 MW; AD3G99E47618FD37 CRC64;

Query Match      24.4%; Score 1182; DB 5; Length 1059;
Best Local Similarity 29.2%; Pred. No. 1.5e-89;
Matches 323; Conservative 145; Mismatches 347; Indels 292; Gaps 38;

9 RWLPFLILLILLPLARAGAPRPAAGQASR-----LVYPTL----- 47
11 RWLG-LLLHWLVVVTQGGVPRPLYGLSHSDLVAGEGLVVRVHPDGAFTMOLEYA 68
48 -----PGSAGELALHLNARFGKGFVRLAPDSDFLAPFEKIERLGGSG 89
69 HELDHRHRHQRSLNSEHDTQAADLHLLPLANETLHEIMHVSFLAPNLVVER----- 123
90 RATGSGRGAR-----GPFSGTYNGEBSLAAYSLGCL----- 123
124 ---HRRDRTSPSLTTRHNLNHFQKVGKRGQATVVAISTGALVSCNFCVLPBILITLC 179
124 -SGSFLDGEFTI-----CPQAGAG-----SLAQPHRLQKRWGAGA-- 159
180 QVGHITATANEYFIEBSKHEHPNQHVVFGRSSVYKSLKRNKRKRGKSGSGA 239
160 -----RPLRPGPWE-----VETSGQQRQEGDHOEDDEBSQDEEAGAGASP 202
240 EVSNCGTRPRRMRMETRLQWQKGVKQVGGQIRHHHHHHHHKRYRHQKISLV 299
203 P-----PPLGATSRKRFVSEARFVETILVADSMAAFYGDLONHITLIMS 249
300 PHTKFYETQPTQEDDHAIEPRRRSISPRVETLIVADAMSAFH-RDLNGYLLITLN 358
250 VAAATYKPSISKNSINIMVVKYLVIDEKMGPE--VSDNGCLITLNFQWQRRENPQSDR 307

```

Db	359	WSALYEDPSIGNSIEIVVRIITQIDEEBSQCLNLTNOAQNIDRFCSOMHKNKSEK	418
Oy	308	HPEHYDPAILLTQONFCQCGGLDITLVADIDITIDDPKSCSVIEDEBLQAAHTLAEIG	367
Db	419	DPEHHDAIILTRKNICDANN - OMTGLANVGMCKPKYQSCSWEDNIGIMLSTHTEIG	476
Oy	368	HVLSMPEDSK - PCTRLFGPMGRHVAFLFYHNLQTLPWSECSAMYITELLDGHDCL	426
Db	477	HNFMFEDTAKIGCHPRVGBI - VHMTPDTFGADITQYOWMSCKSKYITTHLDGQLECL	534
Oy	427	LDAPGALPPE --- TG - LFGNALYQLDQCRQIFG - PDRHCENTSAODVAQJWCH	479
Db	535	DDPE --- TPLDEYNVYGEPLFGMR - VNAQGQRLFNITLTSBVGACSAPEFSTLMCK	569
Oy	480	TDAEPILCHTKNSLPMADGTPCGPGLTSESCCP.EEEVERPKVVDGMAFGPMGEC	539
Db	590	VNG - ECTYHNR --- PLAPGTLCKRKNQNGKCYRREBL --- AAVNGGDMSEK	640
Oy	540	SRTCGGVQFSHRECKDPEPONGARYCLGRRAKYOSCHTEECPPDGKSFREOQCEKNAY	599
Db	641	SRSCGGVSVTQORECONPVPANGVYFCIGERKRYKICRKRPCPEEBSFFRAQCCARDNV	700
Oy	600	NYTDMOGLLOWVPKXAGVSPRRCCLFCRARGSEPFVEFKYIDGTLGSEPTLACVR	659
Db	701	SY --- OGATYKMLPFEDKXNP --- CKLFCSDYDITIANMGATYDGTICTLGTNNACID	754
Oy	660	GQCVKACGDHVDSFKLDCXGCVGCKGNSCKRVSGSLT --- PTNYGNDIVITIPAGAT	715
Db	755	GICRKCVCMDIVDSEVQDDRCGCGSGSDQCQPVARETDTDPAAADGAYVEIVTIPAR	814
Oy	716	NIDVKQRSHGQVNDGNYALTKTAQD - QYLLNGMALISALEODLVKGTILKXSGSIAT	773
Db	815	HILIRE --- LANSHPFLAIAGDGDGFYLNIGSLISMGBE --- FELAGAESL	861
Oy	774	LESL - OSFAPLPEPL - TVOLLT - VPGEVFPFKYTFYFVNDVDFSMQSKERATTNI	828
Db	862	YDRVDEBEITTIQPIQHSISLAIYRGNESNAGIFYEFTLP --- ALANTYAR	911
Oy	829	IQPLIHAQVYLGWSECSSTCGAG ---	852
Db	912	--- QPOMRLSNWTKCSACSGGVQHREPIQENGKESNEPFOIVSIKNNLILAGD	966
Oy	853	-----WQ-----RTVBCRD	862
Db	967	TLPWTAKKRPARQSRGCGDPCPCAHMPPGMPQFCVTCRPGVFAVAPPRKRSVCLD	102
Oy	863	PSG - QASATCNALXPEDAKPCESOL	887
Db	1027	EHYVVAADAECHGLQKPAEMEPRESSL	1053
RESULT 10			
Q08BKT1 PRELIMINARY; FRT; 1009 AA.			
AC	Q08BKY1		
DB	01-MAR-2003 (Tremblrel. 23, Created)		
DT	01-MAR-2003 (Tremblrel. 23, Last sequence update)		
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)		
DE	ADAMTS-12 precursor (Fragment).		
GN	AI605170.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
RA	NCBI_TaxId=10090;		
RE	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Head;		
RX	MEDLINE=22354683; PubMed=12466851;		
RA	The PANTON Consortium.		
RT	the RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
	60,770 full-length cDNAs."		

RL Nature 420:563-573(2002).  
 DR EMBL, AK048612, BAC33391.1, -.  
 DR PIR, P70546, P70698.  
 DR MGI, MGI:2146046, A1605170.  
 CG, GO:0004422, F:metal ion binding, IEA.  
 DR GO, GO:0008270, F:zinc ion binding, IEA.  
 DR GO, GO:0005608, P:proteolysis and peptidolysis, IEA.  
 DR InterPro, IPR001560, Peptidase\_M25.  
 DR InterPro, IPR002870, Peptidase\_M25\_N.  
 DR InterPro, IPR006025, Pept\_M\_zn\_BS.  
 DR InterPro, IPR000884, TSP1.  
 DR InterPro, IPR008085, TSP\_1.  
 Pfam, PF01562, Pep\_M25\_pripep, 1.  
 DR Pfam, PF01421, Reprolysin, 1.  
 DR Pfam, PF00090, tsp 1, 4.  
 DR PRINTS, PRO1705, TSP1REPEAT.  
 SMART, SM00309, TSP1\_4.  
 DR PROSITE, PSC0215, ADAM\_MERO, 1.  
 DR PROSITE, PSC0092, TSP1\_3, 1.  
 DR PROSITE, PS00142, ZINC\_PROTEASE, 1.  
 DR NOD, IER\_1005, 1009  
 SEQUENCE 1005 AA, 113352 MW, 7B63218CFE0FD81 CRC64

Query Match 23.9%; Score 1161.5; DB 11; Length 1009;  
Best Local Similarity -32.5%; Pred. No. 7.4e-88;  
Matches 287; Conservative 159; Mismatches 314; Indels 123; Gaps 37

[illegible]

```

Db      693 DRGVCIGDGSACQTVKKLFRQKGS-----GYVDILIRKARDIRWME-----IKXA 744
Qy      731 GNYLAKTADGQ-YLNNGNLALSAIEQILYKGTILKXSSGATLRLQSRPLPEPLTV 789
Db      742 GNFALRISSEDEKRYLNGGFTIQ-WNGVYKLAGIVFOYDRK-GDLEKILAGFTNESVWL 799
Qy      790 QLTIVPEVPEPKVYETFFVENDVDFSMQSKERATTNIIQPLHMAQVWLDWSECSTC 849
Db      800 QLLP-----QVTPGAIKEYETVRKD-----GLNDVDEKLYTF-WQGRMTCSGYTC 844
Qy      850 GAGWQRRTVEC-RDPSGQASAT-CKALKPE-DAPCPSQICP 889
Db      845 GTGIRQAAHCVKRGHGIVKTTFCNPETQPSVRQKGEKDCOP 887

RESULT 11
G811B3 ID O811B3 PRELIMINARY: PRT: 1600 AA.
AC O811B3
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Metalloprotease disintegrin 12 protein.
GN ADAMTS-12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
CX [1]
RN SEQUENCE FROM N.A.
RP Cal S., Lopez-Otin C.;
RA "Mouse ADAMTS-12,"
RT Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL: AJ537452; CAD50967.1; -.
DR PIR: P05946; P0698.
DR GO: GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO: GO:0008270; F:zinc ion binding; IEA.
DR GO: GO:0007223; P:integrin-mediated signaling pathway; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR001590; Peptidase M12B_N.
DR InterPro: IPR002870; Peptidase M12B_N.
DR InterPro: IPR006025; Pept_M_zn_bs.
DR InterPro: IPR008084; TSP1.
DR InterPro: IPR008085; TSP 1.
DR Pfam: PF01562; Pfam_M12B_propep; 1.
DR Pfam: PF01421; Reprolysin; 1.
DR Pfam: PF00090; Tsp 1; 7.
DR PRINTS: PR01705; TSP1REPEAT.
DR SMART: SM00209; TSP1; 8.
DR PROSITE: PS00215; ADAM_MEPRO; 1.
DR PROSITE: PS00092; TSP1_6.
DR PROSITE: PS00142; ZINC_PROTEASIS; 1.
KW Integrin; Protease.
SQ SEQUENCE 1600 AA; 177791 MW; 083089D356B38C5D CRC64;

Query Match 23.9%; Score 1161.5; DB 11; Length 1600;
Best Local Similarity 32.5%; Pred. No. 1,5e-87;
Matches 287; Conservative 159; Mismatches 314; Indels 123; Gaps 37

Qy      49 GSAGELALHLASFGKGFVILRLAPDDSFAPAEKIERLQSGR-----ATNGERGILRGCF 103
Db      86 GSGDDLYRISHEEKDLFNLTVNMEFLSNGYVERKRYNLSHVKNVASSGQ---PCHL 141
Qy      104 SGTVNGEPEPSL-----AAVSLCRGLSGSFLLLDEEFTIQ-----QGAGSLAQPHRL 151
Db      142 RGTIVLQGGTVIGIGFNALASACQGLTGFHLPGHDFIEIVKKGHLTEE-----GYPHV 196
Qy      153 QMWGAGARPLRGRPEWEVETEBGQROEKGHDQEDSEESOESEEAAGASEPPPGATSR 211
Db      197 YR-----RQSTRAPF-----TKEPICGLKSDLSNVKQELQREKREKTLRSRL 241
Qy      212 TRKPFSEARFVETLLVADASMAAFGAD-LQNHILTLMSVAARIYKPSIKNSINIMLVK 270

```

242	SRSISIKEMWVETLLVADDTKIVEYHSGENVSAYLLITMMNWTGLFHSPTSIGNLHWIVR	301D
QY	271 VLLIVEBKMGPEVSNQGLTLRNFCNWRERNQSDSHPEHYDPAILLTQNFQ-CGOEGL	359D
Db	302 LILLIEEBEGGLKIVHAEKTLSSFFKQKNSINPSDLNPNVHVAVALITTKDICAGNRP	361D
QY	330 CDTLVADIGTTCDPNKSCSVTEDEGLQAATLAHLAIGHVLSNPHDSK-PCTRLFGPMG	368D
Db	362 CERTGLSOLSGMCOPHRSGCNINBDSGLPFLAFTIAHEHGSFGIHDKEKDCE---FVG	417D
QY	389 KH-HWNAFLFVHLNQTLFPMSPCSAMYLLELLDDGHGDCDLDAFGAALPLPTGL-----P	441D
Db	418 RHRYTNSQOIQYDPPRLPTMSKSKNYITRFLDRGRGLDD-----IPSKGGLKSNVIAP	472D
QY	442 GRNALYQLDQCRQIFGDPFRHCNTSAODYCAOLMCHTDGAEP LCHTKNGSLPMADGP	501D
Db	473 G--VIVDVHHCQOLOGPNATPC--QEVENVCQTLWCSSVKK---FCRSKLDA--ADGTR	523D
QY	502 CGPGLHCSSGSCLPSESEVERPKPVVDGMAFMGPMGECSPFCGGGVFHSRECKDPEPCN	561D
Db	524 CGEKKCMAGKCI---TYCKRPSIIPGMGMNPSFMHSHSRICGAGAGASABRLCNPPPKF	560D
QY	562 GGRYCLGRPAKIQSCHTEBCPPDGKSFEEQCEKNAVNTYTDMDGNLLQNVPKAGVSPR	621D
Db	581 GGRYCGEGRKRYRLCNVHPCREDTPTFOMOCSEBDIVFYKN---QFYRPFVFNAAHP-	639D
QY	622 DRCKLFCRAGR--SEFKVEAKVVDGLTC--GPETLAI CVRQGVACGCDHVVDSPRL	677D
Db	637 --CELYCRPIIDQFSR-RYLEH-VLDGIPCEGGNSRVVCINGI CKRGVGDYEIDSNATE	692D
QY	678 DKQGVCGGKGNK-----KRVSSSLPTNYGNDVITPAGNTINDYKQSRHGVND	730D
Db	693 DRGCVLIGDSASACQYVKLFRQKES-----GYVDIGLIPKAGDIRIVE---IKAA	741D
QY	731 GNTLAKTKTDGQ-YLLNGNLATSAIEODIILVKTILTKSGSIATLERLOSFRPLPELTV	789D
Db	742 GNTLARSDEPEKTYNGFPIIQ-WNGVYKLAGTVFOYDR-GPLEKLIAPGNFESVWL	799D
QY	790 QLLTVBGEVPPPKYKTFVENVDPVDSMQSSKERATNTYIOPLLHAQWVLDGMEGSCSTC	849D
Db	800 QLLF--QVYNPQIKXEYTVRD-----GLNDVBEKLLLYF-WQFGWTECSVTC	844D
QY	850 GAGMORRTVEC-RDPSGQASAT-CNKKALPE-DAPPCESQLCP	889D
Db	845 GTGIRGQAHCYKKGIGYIKTFCNPNENQPSVROKCKHEKQCP	887D
RESULT 12		
Q8CG28		
ID	Q8CG28	PRELIMINARY; PRT; 1070 AA.
AC	Q8CG28;	
DT	01-MAR-2003 (TREMBLrel. 23, Created)	
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)	
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)	
OS	Zinc metalloendopeptidase.	
OS	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
NCBI_Taxid=10090;		
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=129/Sv.	
RA	Brachnate M., Nagaraja R., Abe K.;	
RT	"Genomic Sequence Analysis in the Mouse t-complex Region.";	
RL	Submitted (JUL-2002) to the EMBL/Genbank/DBD databases.	
DR	EMBL; AF528163; AAC17380.1; -	
DR	GO; GO:0004222; F:metalloendopeptidase activity; IEA.	
DR	GO; GO:0008270; F:zinc ion binding; IEA.	
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.	
DR	InterPro; IPR001590; Peptidase M12.	
DR	InterPro; IPR002870; Peptidase M12B_N.	
DR	InterPro; IPR006025; Pept_M_Zn_BS.	
DR	InterPro; IPR000884; TSFL.	





OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 ON NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkley.  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
 RA Miranda A., Mungai C.J., Nuno J., Pacleb J., Paragas V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Ceoliker S.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY119538; XAM50192.1; -.  
 DR FLYbase; FBgn0029791; CG4096.  
 DR GO; GO:0004222; F:metallopeptidase activity; IEA.  
 DR GO; GO:0008270; F:zinc ion binding; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001590; peptidase\_M12B.  
 DR InterPro; IPR002870; peptidase\_M12B\_N.  
 DR InterPro; IPR006025; pept\_M12B\_BS.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008085; TSP1.  
 DR Pfam; PF01562; Pep\_M12B\_propep; 1.  
 DR Pfam; PF01421; Repolysin; 1.  
 DR Pfam; PF00090; tsp\_1; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00209; TSP1; 1.  
 DR PROSITE; PS50215; ADAM\_MEPRO; 1.  
 DR PROSITE; PS50092; TSP1; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 SQ SEQUENCE 769 AA; 86543 MW; 39404FBD271E1968 CRC64;

Query Match 20.1%; Score 976; DB 5; Length 769;  
 Best Local Similarity 31.5%; Pred. No. 1,76-72;  
 Matches 251; Conservative 103; Mismatches 279; Indels 164; Gaps 26;

QY 9 RWLPFLLLLLLLFLARAPAPPAAGQAS-----LVVPTRL----- 47  
 DB 11 RWLG--LLHFWLVVVGQGVPRLYGSHDELVAEGQLVPRRVHPDGAFTHQLEYA 68  
 QY 48 -----PSAGELALHLSAFGKGFVLRAPDSTFLAPFKTERIGSG 89  
 DB 69 HELDHRHRORRSINSEHDTQAADLHLPLANETLLELMASYLEAPLVYER----- 123  
 QY 90 RATGGERGLR-----GCFPSGTVNGEPESLAVSLCRGLSGFLDGEFTI--- 136  
 DB 124 ----HRDRLTRBPLTRRLNCHFGKVRGQPAITVAISTCAGLVGHIRTAGNEYFIEPS 179  
 QY 137 ---OPQAGG-----SLAOPRLQWGPAGA-----RPLPRG 165  
 DB 180 KEHHPHVNHPVVFQRSSVKKPKSHLRKKNKRGKSGSGAEVSNCGTREPFRMRMR 239  
 QY 166 PEME-----VTGEGORQORGDHQDESESEGEAEAGSEPP-----PPL 206  
 DB 240 LEWQARQKVVQGRQIRRHNNHHNNHGHKTRHHQOKISRVPTKFKYETQFQTEPDH 299  
 QY 207 GATSRTRFVSEARFVETLLVADASMAAFYADLQNHILTLMSVAARIYKPSIKSINL 266  
 DB 300 AEIPRRRSISSPRHVTLLVADATWSAFH-RDLNGYLLTIMVSAIYKPSIGNSIEI 358  
 QY 267 MYKVLIVDEKGRPE--VSDNGGLLRNCWQRFNPSDRKPHHYDTAILITRQNF 324  
 DB 359 VVVRITQLDEESQLQNTLQNAQKLDRCFSQWQHLKNGSEKDPHHHVAAILITRKNIC 418  
 QY 325 GQEGLCITLGVADIGITCDPNKSCSVLEDEGLQAHTLAHELGLVLSMPHDSK--PCTRL 383  
 DB 419 ANN--CWTGLAVAGCKRCKKQSCSVNEDNGIMLSHTITHELGNFMGHDTAKIGCHPR 476  
 QY 384 FGPMGKHVVAFLFVHLNQLTLPWSPCSAMVLTLLDGGHDCLLDAPGALPLP---TG 439

DB 477 VGPI--VAIMPTREADTLQVCMNSCRKTIITHLDOGLGECIDDP-----TPLDEYNTG 531  
 QY 440 -LPGEMALYQLDQCRQIFG--PDRHCPTNSADVCAQMLCHTDGAEPICHTKNGSLPW 496  
 DB 532 ELPGWR--YNARGQCRQLFNLITDSEVAGCAFPHEFGSLTMCXNG--ECVTHMR---PT 584  
 QY 437 ADGTPCGGHLCSBSCCLPEEVEPRKPVYDGNAPNPMGECGRTGGGVQFHSRECKD 556  
 DB 585 APGLCGNKNKCQNGKCVREEL---AAVNGWGDSEWSESCSRSCGGSVSTQGRBCDN 640  
 QY 557 PEPONGRYCIGRRAKYQSCHEECPPDGKFRQCEKXNAVYTDGNNLQWVPKXA 616  
 DB 641 PVPANGGVFCIGERKRYKICRKCPCPAEPSPFRAQGCARFNVS---QGATYKMLPFED 697  
 QY 617 GVSPPDRCKLFCRARGSEFVFEAKYIDGTLGPEPTLALCVRQCCKVAGCDHYVDSBRK 676  
 DB 698 KNNP---CKLFCSDVDITLIANMGATVLDGTPCTLGTNNMCIDGIC----- 740  
 QY 677 LDKGVCCKGKNSCRKV 693  
 DB 741 --KLVAIGSSTRRCRTI 755

Search completed: May 7, 2004, 11:51:07  
 Job time : 60 secs